UNIVERSITY OF STELLENBOSCH

THE ANALYSIS AND QUANTIFICATION OF UNCERTAINTY FOR LEAST LIFE-COST ELECTRICAL LOW VOLTAGE DISTRIBUTION DESIGN

Schalk W Heunis

Dissertation presented for the
Degree of Doctor of Philosophy in Engineering at
The University of Stellenbosch

Promotor:
Prof Ron Herman, University of Stellenbosch

December 2000

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree

Date: 4 December 2000

Synopsis

The purpose of this dissertation is to provide methods for designing and managing low voltage residential feeders. These methods can be applied to the problem of planning residential networks under uncertainty while ensuring least life-cycle costs. By analysing collected load data from various communities in South Africa, a new probabilistic model for representing the load uncertainty of residential consumers was derived.

This model uses the beta probability distribution to describe individual consumer loads over a period of time. Methods for combining the loads in linear combinations were used to derive a new probabilistic voltage regulation calculation procedure. This new method is different from previously developed voltage calculation methods in that it can be used to estimate the probable voltage performance of a feeder over a period of time. A simplification of the method is proposed which allows it to be implemented in any commercial spreadsheet program.

The new probabilistic load model was also applied to the problem of calculating resistive losses in residential low voltage feeders. A new probabilistic method was formulated and this method can be used to estimate the probable range of resistive loss in a feeder for a period of time. This method is simple enough to implement in a commercial spreadsheet program.

Probabilistic information about network and load parameter uncertainty is seldom available and these uncertainties are best modelled using fuzzy numbers. The probabilistic calculation methods cannot represent these uncertainties and only after applying a fuzzy-probabilistic approach can both types of uncertainties be used. This is a significant enhancement to the current methods and ensures that the uncertainty about the calculated results is realistically represented.

The specification of load parameters for the methods was significantly simplified following a regression analysis of collected load data from South African communities. By specifying the distribution of the consumption of individual consumers in a community, the other load parameters can be estimated using a set of fitted linear regression equations. This greatly

reduces the burden of specifying the load parameters and makes it possible for the proposed calculation methods to be applied to the design of new feeders in practice.

The distribution of the consumption of individual consumers can be specified using the average and the standard deviation of the consumptions of individual consumers. Accurate estimates of these parameters can be obtained from sales information and can be used to manage existing networks effectively. Using the sales information with the proposed methods enables more cost-effective upgrades of existing feeders low voltage feeders. The identification of potential problems in existing low voltage networks is also possible if the layout of the feeders in a community is known.

The use of the proposed methods is illustrated in step-by-step fashion. Typical input parameters are used and all the required calculations with intermediate results are presented.

Sinopsis

Die doel van hierdie proefskrif is die daarstelling van residensiële laagspanningsnetwerk ontwerp- en bestuursmetodes. Hierdie metodes kan toegepas word vir die beplanning van residensiële laagspanningsnetwerke waar onsekerheid bestaan oor toekomstige kragverbruik en die spesifikasie van die netwerkparameters. Lasdata, wat versamel is in verskeie Suid Afrikaanse gemeenskappe, is geanaliseer en 'n nuwe probabilistiese modellering van die onsekerheid oor die kragverbruik van residensiële verbruikers is ontwikkel.

Gebruik is gemaak van die beta waarskynlikheidsdightheidsfunksie om die tydsgebonde kragverbruik van die verbruikers voor te stel. 'n Nuwe probabilistiese spanningsvalberekeningsmetode is ontwikkel en die metode maak gebruik van liniêre kombinasies van die lasstrome van die verbruikers. Die verskil tussen hierdie metode en bestaande metodes is dat dit die tydsgebonde waarskynlikheid van die spanningsregulasie van 'n kabel kan bereken. 'n Vereenvoudiging van die metode is ook verkry en dit kan in enige kommersiële sigblad geïmplementeer word.

Die probabilistiese lasstroommodel is ook gebruik om 'n nuwe probabilistiese energieverliesberekeningsmetode te ontwikkel. Hierdie metode kan gebruik word om die tydsgebonde waarskynlikhede van 'n reeks van moontlike energieverlieswaardes te bereken. Die metode is eenvoudig genoeg om in enige kommersiële sigblad te implementeer.

Onsekerheid oor die spesifikasie van die parameters van die nuwe metodes asook die netwerkparameters kan nie met probabilistiese metodes voorgestel word nie, aangesien inligting oor die waarskynlikhede van parameters selde beskikbaar is. Hierdie onsekerhede kan beter voorgestel word deur die gebruik van sogenaamde "fuzzy"-metodes. Die voorgestelde probabilistiese metodes is aangepas om hierdie tipe onsekerhede ook in ag te neem. "Fuzzy-probabilistic" metodes is gebruik vir dié aanpassings en word beskou as 'n noemenswaardige verbetering van die metodes. Die verbeterde metodes verkaf meer realistiese voorstellings van die onsekerheid oor berekende resultate.

‘n Statisitiese analise van Suid Afrikaanse lasdata het ‘n vereenvoudiging van die spesifisering van die parameters van die nuwe metodes tot gevolg gehad. Die waarskynlikheidsverspreiding van die energieverbruik van huishoudelike verbruikers kan gebruik word om akkurate skattings van die ander parameters te verkry. Hierdie vereenvoudiging het tot gevolg dat die nuwe metodes vir praktiese netwerkontwerp gebruik kan word.

Die waarskynlikheidsverpreiding van die energieverbruik van verbuikers is beskikbaar in die vorm van energieverkope en kan gebruik word vir die effektiewe bestuur en opgradering van bestaande netwerke. As die uitleg van die bestaande netwerke in ‘n gemeenskap beskikbaar is, kan die inligting wat bevat is in die energieverkope gebruik word om probleme in bestaande netwerke te identifiseer.

Al die voorgestelde metodes is stap vir stap uiteengesit met voorbeelde van al die berekeninge met tipiese waardes.

Acknowledgements

I would like to thank the following people without whom this thesis would not have been possible.

- My supervisor, Dr Ron Herman, for his support and encouragement
- Marcus Dekenah and the NRS Load Research project for supplying me with collected load data and several great ideas
- Eskom, for financial support
- My wife, Louise for her love and encouragement
- My Creator, for allowing me to have a bit of fun

This dissertation is dedicated to Louise

Table of contents

List of symbols

a_i	is a constant associated with load current i and is used to calculate the voltage regulation based on the principle of superposition
α_k,β_k	are the parameters of the beta pdf of the load current for consumer k
α,β	is the parameters for the beta distribution function
C_k	is the circuit breaker size or maximum value of the load current for consumer k
Cl	is the confidence level (pu)
Co	is the intercept of the linear regression between δ_k and σ_k
C_{o2}	is the intercept of the linear regression between δ_k and σ_k^2
δ_k	is the difference between μ_k and $E(\mu_k)$ for consumer k
$E(\rho)$	is the expected value of the correlation between all the consumers in a group
$E(\mu_k)$	is the expected value of the average current traces for a group of consumers, calculated for a period of time
$E(\delta^2)$	is the variance of the individual consumer current means and is equal to $V\delta$
$E(\rho_{ij})$	is the expected value of the correlation coefficient and is assumed independent of the standard deviations of the load currents of consumers i and j.
$E[\delta^3]$	is the third moment of the individual load current means
G	is the slope of the linear regression between δ_k and σ_k
G_2	is the slope of the linear regression between δ_k and σ_k^2
Maximum	is the upper bound of the scaling range for μ_{ns} and σ_{ns}
Minimum	is the lower bound of the scaling range for μ_{ns} and σ_{ns}
μ	is the average of the individual load current means calculated for a period of time
μ_k	is the mean of the current for consumer k, calculated for a period of time
μ_{loss}	is the average loss for a period of time
μ_{ns}	is μ_k scaled to the range 0 to 1
Qa	is a matrix where element k in column s, is 1 if consumer k is supplied by phase A in section s else the element is 0
Qb	is a matrix where element k in column s, is 1 if consumer k is supplied by phase B in section s else the element is 0
Qbeta	is the inverse beta function
Qc	is a matrix where element k in column s, is 1 if consumer k is supplied by phase C

	in section s else the element is 0
ρ_{ij}	is the correlation between load current i and load current j
$\rho_{\delta\sigma}$	is the correlation between the individual consumer load current averages and standard deviations
$\rho_{\delta\sigma^2}$	is the correlation between the individual consumer load current averages and variances
Rn_i	is the neutral resistance from the source to consumer i
Rns	is the neutral resistance of the cable in section s
Rp_i	is the phase resistance from the source to consumer i
Rps	is the phase resistance of the cable in section s
σ_{ns}	is σ_{ns} scaled to the range 0 to 1
$T1_k^2$	is the value of the first derivative of the average loss to δ_k with δ_k=0
$T2_k^2$	is the value of the second derivative of the average loss to δ_k with δ_k=0
$V[\mu_{loss}]$	is the variance of the average loss
σ^2_k	is the variance of the current for consumer k, calculated for a period of time
$V\delta$	is the variance of δ_k in a group of consumers
$V\delta^2$	is the variance of the squares of the individual load current means
Vs	is the supply voltage (230 V in South Africa)
$V\sigma$	is the variance of the individual consumer load current standard deviations
$V\sigma^2$	is the variance of the individual consumer load current variances
z_α	is the percentile value of the standard normal distribution at some level of confidence, α

1. INTRODUCTION

1.1 BACKGROUND

In South Africa in 1992, the newly elected government promised "electricity for all" and expected the largely state-owned utility, Eskom to meet this promise. Eskom invested capital in the expansion and construction of new medium (22 or 11 kV) and low voltage networks (400V). By the end of 1999 more than 1.5 million new homes had been supplied with an electrical connection.

Eskom and the South African government are still (in the year 2000) committed to providing electrical connections to the households not yet connected. A large number of future connections will be provided to rural households that are located far from the existing infrastructure. This means that the cost per connection will be very high and designers will be faced with the challenge to reduce these costs.

The management and upgrading of existing networks is a further challenge planners will face in future. In South Africa the distribution industry is headed for restructuring and the new regional distributors will have to optimally utilize the existing infrastructure. The existing infrastructure was constructed at a time when little was known about residential consumers and their load. A number of assumptions were made about how the load of consumers will grow and what the demand of the mature networks will be. Unfortunately not all of these assumptions were correct and it meant that some of the networks will never be fully utilized, while others will have to be stretched to meet the consumers' demand.

Other developing countries face similar challenges and could benefit from the experience gained by South African researchers. In South Africa a number of design tools have been developed using the results of load research projects. Similar load research projects have been initiated in neighbouring countries and it is envisaged that the developed tools could be used and calibrated for use in these countries.

The work presented in this thesis was initiated with the aim of providing utilities in developing countries with better tools for planning residential low voltage feeders under uncertainty.

1.2 DESCRIPTION OF THE PROBLEM

Planners of residential feeders in South Africa face two problems:

- How to design new low voltage feeders under uncertainty to ensure least life-cycle costs
- How to cost effectively upgrade existing networks while ensuring that the life-cycle costs of the feeders are minimized

The life-cycle costs of residential low voltage feeders comprise the following

- the initial capital investment,
- the depreciation of the invested capital,
- the cost due to resistive losses and
- any additional capital investments to ensure that the quality of supply is acceptable.

Due to the type of conductor used in South Africa to construct the low cost networks and the length of the typical feeders, the primary constraint in the design of the residential low voltage networks is the voltage regulation on the network.

Some of the existing residential networks in South Africa were designed so that they can easily be upgraded after a number of years (typical 7 years). The ability to easily and cost effectively upgrade networks, hedges some of the risk of planning in an uncertain environment. After a number of years, the planners of the network will also have a better indication of the accuracy of their initial estimates of the mature load of the network. Ideally, the initial network designs will be revised to ensure that the upgraded networks are as cost effective as possible.

Even residential feeders that were designed without an upgrade path might need to be upgraded. If the voltage performance of the networks does not meet the statutory requirements the design of these feeders will have to re-evaluated and adjusted. In this scenario, better estimates of the loads of the consumers will ensure cost effective upgrades.

1.3 CURRENT DESIGN PRATICE IN SOUTH AFRICA AND APPROACH FOLLOWED

The Herman Beta voltage regulation design method is currently the recommended [51] design procedure of low voltage feeders in South Africa. This procedure uses a probabilistic description of the load currents of residential consumers at the moment of system maximum demand.

Many points of high demand however occur during a period time and the Herman Beta method does not account for this. If the voltage regulation on a low voltage feeder is calculated using the Herman Beta method, a range of probable values is obtained. By applying a confidence level to the range of values, a design percentile can be extracted. If, for example, a confidence level of 90% is used, the design percentile can be interpreted as the level of voltage regulation above which 90% of similar feeders will operate at the time of high demand.

Consider, however the same designed feeder at a different period of high demand. If the distribution of the voltage regulation for different feeders at this second period was equivalent to the first period, a 90% confidence level would have the same implication as the first period of high demand. This presents the designer with the following problem: what is the probability that the voltage regulation on the designed feeder will be higher than the design percentile during both periods of high demand?

Instead of two periods of high demand, the problem can be extended to: What is the probability of the voltage regulation on the feeder exceeding the design percentile for a given period of time or during the life-cycle of a residential feeder? Knowledge of this probability, would allow a planner to playoff the reward in reducing the initial capital investment with the risk of doing additional upgrades on the system.

To extend this argument to the life-cycle cost, a probabilistic expression of the resistive losses for a period of time is also required. The planner could then use these probabilistic results to optimise the life-cycle costs of a residential feeder subject to load uncertainty.

The methods proposed in this thesis aim to fulfil these requirements and also provide a mechanism for including uncertainties that cannot be modelled probabilistically due to lack of information. Examples of these uncertainties are load parameter uncertainty and conductor resistance uncertainty. A fuzzy-probabilistic approach provides this functionality and significantly enhances the design techniques.

A large body of residential load and socio-demographic data has been collected by load research projects in South Africa. Analyses of these data led to the identification of strong relationships between some of the load parameters. This reduces the requirements for the design procedures to:

- The average consumption in a community
- The variance of the consumptions of the individual consumers in the community

Consumer consumption is recorded for all residential consumers and this makes the application of the proposed methods ideal for upgrading projects. No additional metering is required to assess communities before upgrades, since the information is already available in billing systems.

The proposed methods were obtained by analysing the load data collected by the load research projects in South Africa. Theoretical modelling of the data and design procedures were tested using Monte Carlo type simulations.

1.4 RESIDENTIAL LOAD RESEARCH IN SOUTH AFRICA

Two residential load research projects are currently (year 2000) collecting load data in South Africa. These two projects are:

- The NRS load research project managed by Marcus Dekenah Consulting [53]
- The TSI load research project managed by TSI and assisted by Marcus Dekenah Consulting [61]

These load research projects are the results of a number of years of research in South Africa. In the late 1980's a team from Stellenbosch University developed a microprocessor-based data logger with the following important features:

- Low cost
- Accurate clock to allow time-synchronized readings to be taken
- Data storage of five minute time stamped average load readings for at least 1 month

A number of these loggers were installed in communities in South Africa and after some pilot studies, Prof Ron Herman of the University of Stellenbosch developed the Herman Beta method, which is currently the recommended design procedure for low voltage feeders in South Africa [51].

This design procedure required a description of the load current of different types of consumers at the time of the system maximum demand. To this end, the mentioned load research projects were initiated in South Africa. The data collection process in an identified community follows a process that can be summarized by the following steps:

- Randomly identify 70-80 households in the community
- Install the microprocessor-based data loggers at the point of supply to the consumer
- Download recorded load data on a monthly basis
- Filter the data and load it into a relational database
- Give feedback to the operators of the data loggers in the field to ensure that loggers that have failed are fixed

During the period of load collection, socio-demographic information is also collected in a front door survey. These data are also filtered and loaded into the relational database. A link is established between the load data and the socio-demographic data for each specific household.

Focus is concentrated on the acquiring of good quality data to ensure that models developed using the data are accurate.

Communities for the load research projects are identified on a yearly basis and the aim is to ensure that representation is obtained in terms of

- Average household income
- Time with electricity

Figure 1 shows an example of a project map that is compiled on a yearly basis to identify areas where load data representation is required. The map in the figure was compiled at the end of 1999 [16] and summarizes the data that were available for this work.

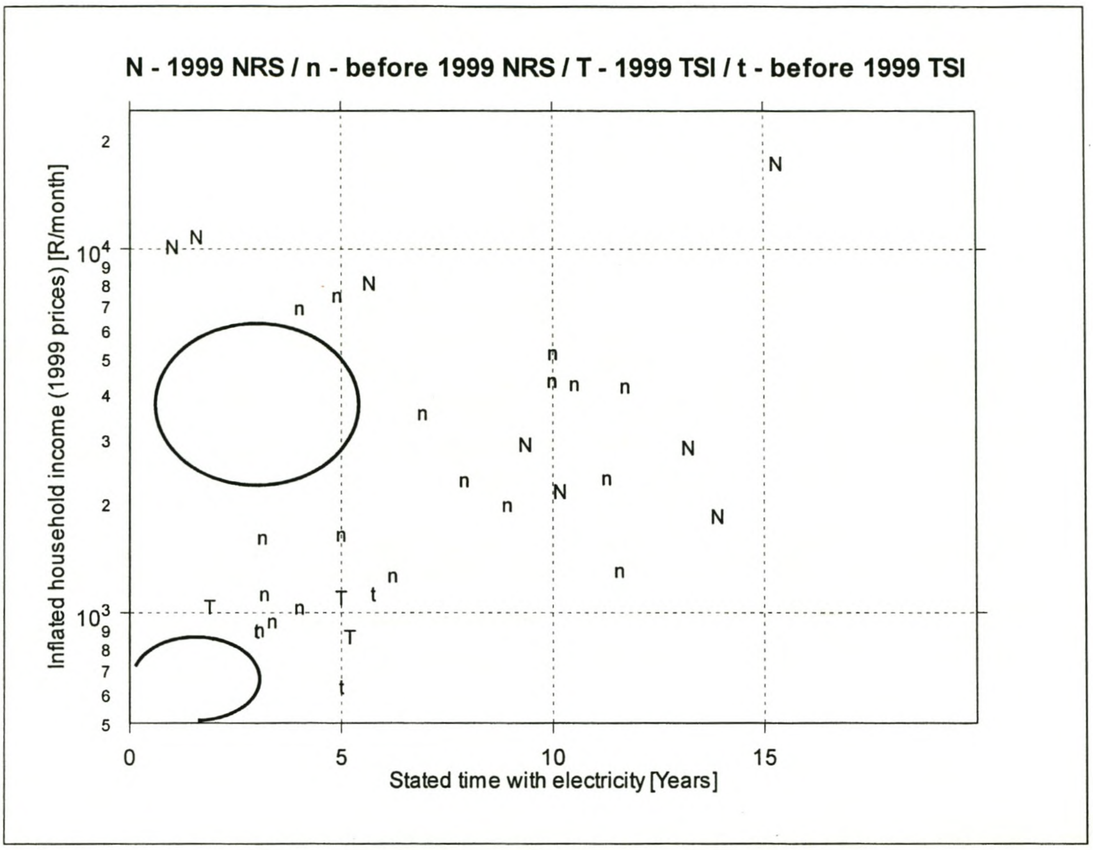

Figure 1 **A typical project map compiled yearly. Regions where load data is required are identified using this map (shown as ellipses)**

More information about residential load research in South Africa can be found in [25]

2. OVERVIEW OF UNCERTAINTIES AND QUANTIFICATION OF UNCERTAINTIES

2.1 INTRODUCTION

This chapter presents an overview of approaches used to quantify uncertainty with respect to published literature. The nature of residential load uncertainty is described and the most suitable approach to quantifying the uncertainty is identified. Load parameter and network parameter uncertainty are considered and the best approach to quantifying these uncertainties is identified.

A method for combining probabilistically modelled uncertainties and fuzzy modelled uncertainties is proposed as an appropriate method for representing the total effect due to the uncertainty of load and network parameters.

2.2 APPROACHES TO QUANTIFICATION OF UNCERTAINTY

2.2.1 Scenarios and weighted scenarios

This approach is perhaps the easiest and most intuitive way to model uncertainty in the absence of any information about the probability of outcome. Some probabilistic information can be included by weighting the scenarios, according to either probability or importance.

The idea is to define a number of possible states for each uncertainty. If probabilistic information or statements about preference are known, then each state can be weighted accordingly [23, 50, 64]. This approach is very useful when evaluating uncertainties with discrete states or uncertainties that can be efficiently modelled with discrete states.

The disadvantage of this technique is that the computational power required to evaluate continuous distributions are enormous. Even with a large number of calculations, the results may not be extremely accurate. A good example of this is where Monte Carlo type simulations are used to represent a continuous distribution of load currents. Although the number of iterations exceeds 1 000 [32] the error can be as large as 16%. The error is calculated using a procedure called bootstrapping.

This model should clearly be used only when a fixed number of discrete scenarios are to be considered.

2.2.2 Probability unknown but bounded

An improvement on modelling with scenarios is a model where the probability distribution of the uncertainty is unknown but the value of the uncertainty can be assumed to lie between two bounds [9, 14, 46].

The range of the uncertainty is limited by defining an upper and lower bound. The advantage over the scenario analysis is that continuous information about the uncertainty can be considered without the knowledge of a specific probability distribution. Interval mathematics can be used to evaluate this type of model [8]. This reduces the computational burden and increases the accuracy.

The disadvantage of this model is that if the upper and lower bounds are not crisp (as is most often the case) or if the upper and lower bounds are estimated too conservatively, the model could lead to significant over- or under-design.

An example of LV feeder over-design is where the circuit breaker limit on a house connection feeder is used as the upper bound of the load current. In most cases, except when the consumers are load limited, this will lead to a significant over design.

2.2.3 Fuzzy set (possibilistic modelling)

When the exact probability distribution of an uncertainty is unknown or is not definable, but more information than just the upper and lower bounds are known, the uncertainty can be modelled using fuzzy sets or possibilistic distributions.

The idea of using linguistic variables or fuzzy set theory was first introduced by Zadeh [65]. An application in long range planning was attempted by Dhar in 1979[18]. Since then an increasing number of publication have appeared in electrical engineering [28, 37, 41, 42, 48, 55,56] and a number of other engineering fields (Earthquake engineering [63]; Hydraulic design [36]; Transportation engineering [60]), including two literature surveys [49 and 58] have been published in electrical engineering literature.

The principle behind fuzzy set theory is that linguistic variables can be used to describe the realisation of uncertainties, for instance system loading is "heavy", maintenance costs are "high", initial load growth is "low". The variables are quantified mathematically as possibilistic distributions [43] or membership functions. A common application of fuzzy set theory is the fuzzy number, which is mathematically described with a trapezoidal membership function. Figure 2 shows an example of a trapezoidal membership function and can be expressed as the quadruplet (a1,a2,a3,a4).

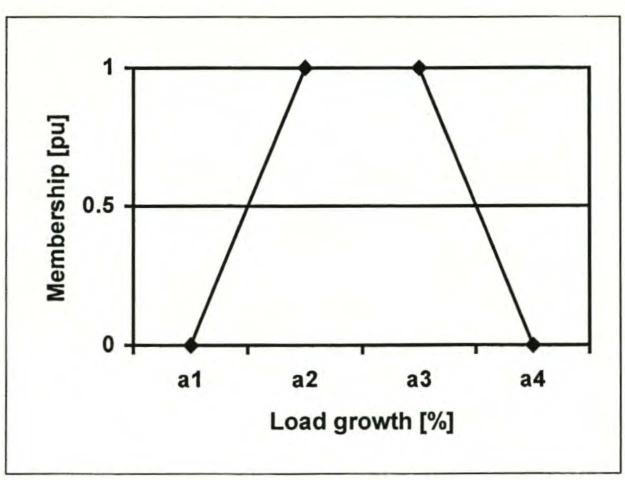

Figure 2 : A membership function of load growth

It should be noted that when a1=a2 and a3=a4, an unknown but bounded representation is obtained (see section 2.2). When a2=a3, a triangular distribution is obtained.

The fuzzy range can be transformed to a crisp range if an α-level cut is made to the distribution. An α-level cut of a fuzzy number includes into a new set, only the values of the distribution where the membership is greater than α ($0 < \alpha < 1$). Since the sides of the triangular and trapezoidal fuzzy numbers are straight, the range can easily be calculated as:

$$\text{Minimum} = a1+(a2-a1)\alpha \tag{1}$$

$$\text{Maximum} = a4+(a3-a4)\alpha \tag{2}$$

More information about uncertainties can be modelled in this way and the bounds for a specific range of values are fuzzy (as defined by the membership function). It is very useful for modelling uncertainties that can be described in human language but are difficult to quantify probabilistically.

The disadvantage of modelling uncertainties with fuzzy set theory is that, in the absence of confidence intervals, unrealistically high or low estimates of the likelihood of an uncertainty can be made.

2.2.4 Probability models

If detailed information about the probability distribution of an uncertainty is known, or can be assumed with reasonable accuracy, the uncertainty can be described probabilistically. Probability models found in literature are:

- Probability distribution functions (pdf) [24,45,59,61]
- Extreme value distribution functions [3,44 p310]
- Modelling of statistical dependence - Markov chains [7,44 373-401], time series models [12,54], multivariate distributions [10,11]

A probability distribution function (pdf) of an uncertainty is a mathematical function that expresses the probability of a specific level of the uncertainty being modelled. A number of different types of pdf exist, for example the normal distribution, the binomial distribution, the gamma distribution and the beta distribution. Each pdf has different characteristics and is used in different applications. For example, the beta distribution significantly fits the domestic consumer load currents at the moment of system maximum demand [31] and forms part of a probabilistic voltage drop calculation method introduced by Herman [33].

An example of a probability distribution function is shown in figure 3.

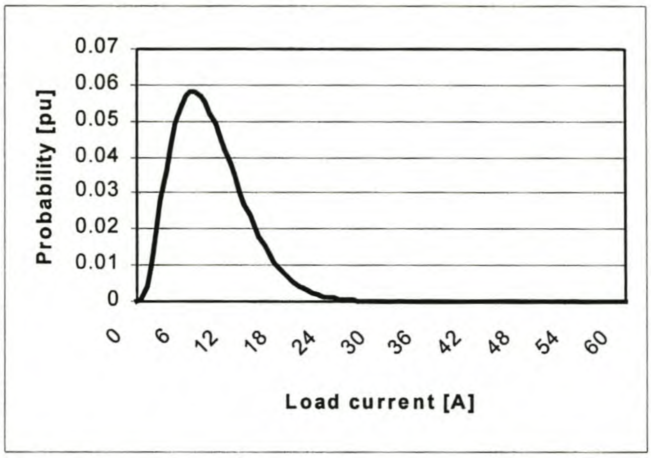

Figure 3 : An example of a load current probability distribution function (pdf)

The *extreme value distribution* is a special pdf, which expresses the probability that a specific level of an uncertainty is the extreme (maximum or minimum) level. An example is found in peak load prediction [3].

Markov chains model time-dependent probabilities, i.e. probabilities that vary with time. The representation of the uncertainty in this case is probabilistic and is dependent on the current state of the system. A transitional probability exists that expresses the probability that the system will change state. An example is found in unit commitment risk evaluation where the energy demand is modelled as a Markov process [66].

Time series models are used to model uncertainties that are highly dependent on previous realisations of the same uncertainty. Examples of this method are found in short term load forecasting[12] and earthquake engineering[54].

Multivariate distributions are used to model more than one statistically dependent uncertainty. When the probability distribution of one variable is a function of one or more other variables, the variables are statistically dependent. An example of an application of multivariate models using a Monte Carlo type technique can be found in [4,10].

The advantage of probabilistic modelling is that statements about confidence intervals can be made if the uncertainties are modelled accurately. Modelling accuracy is a great concern and probability models derived in the absence of sufficient information could give misleading results.

2.3 RESIDENTIAL LOADS

2.3.1 Uncertainty in load modelling

Domestic loads: The electricity demand of residential consumers caused by the use of household appliances.

By analysing this statement, one readily concludes that predicting domestic loads would be almost impossible unless perfect information about the use and ownership or the habits of the consumers are available. Figure 4, a histogram of the energy use in Tafelsig, a medium income community in South Africa, was compiled to show the great variation of consumers within a community [53]. The histogram is further conditioned on geyser ownership that appears to be higher for consumers with a higher monthly income.

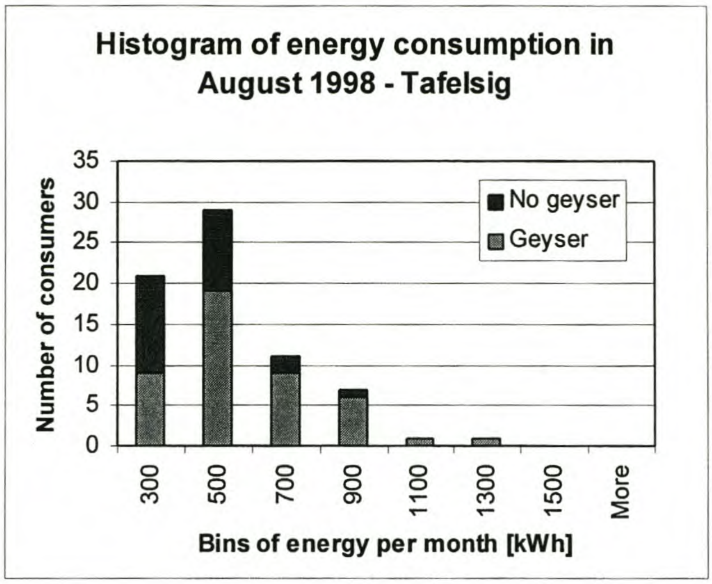

Figure 4 Histogram of energy consumption [kWh/month] as measured in Tafelsig 1998, showing the difference in consumption between consumers with and without geysers

Differences in the use and ownership of appliances are the grounds for load uncertainty. Some factors, which influence the domestic loads, are:

- When connected to an electricity source, appliances might appear to use a **constant current**, regardless of the voltage applied or a **constant power** or might act as a **constant resistance**
- **Temperature** has an effect on the operation and use of the appliances
- Appliances could be used for different **purposes**, heating, cooling, lighting, food processing etc. Differences in the preference of use of appliances and consumer habits lead to uncertainty. Consumers also do not use the same appliances at the same time (**coincidence**).
- Some appliances can only be used when connected to a sufficient **water supply**, e.g. geysers, washing machines etc.
- Appliance penetration into a community changes with time and varies with income, causing **load growth**.

2.3.2 Constant power, constant current and constant resistance

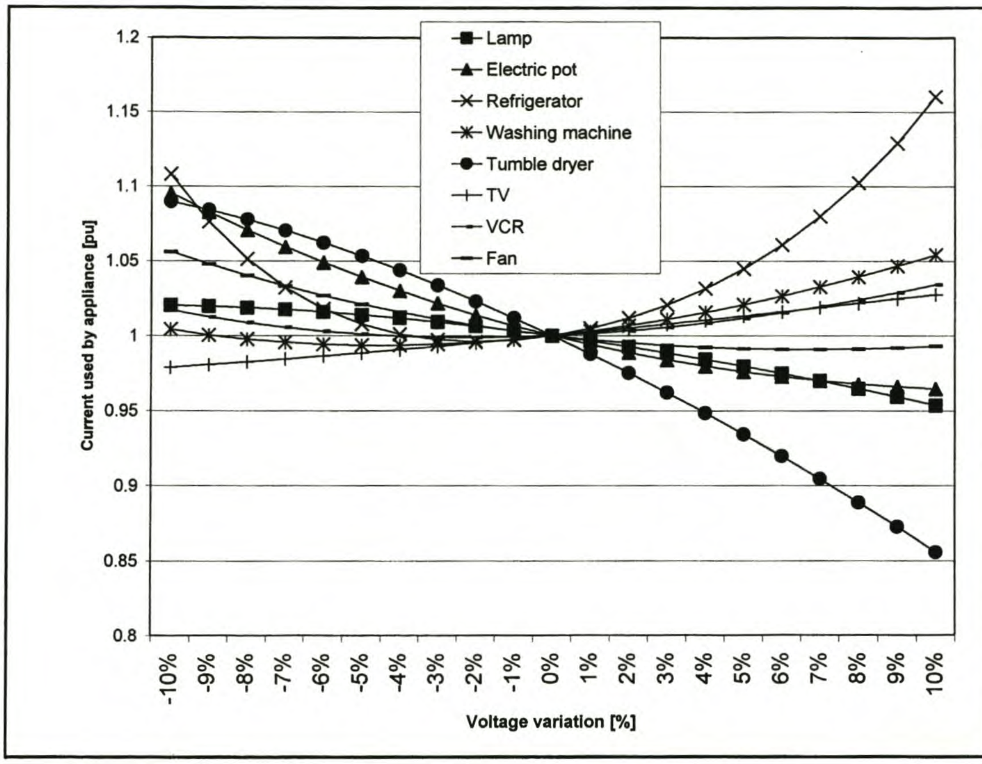

Figure 5 Response of appliances to applied voltage [13]

Figure 5 shows the voltage response of some selected household appliances. The response to voltage, however, also has an effect on the performance of the appliances, which in turn might cause some feedback from the user of the appliance.

The electrical response of an appliance, together with the user's response to the performance of the appliance, causes the appliance to appear as a constant power sink, a constant current sink or a resistor. The aggregated domestic load is therefore a function of the appliances owned by the consumer.

Despite these uncertainties, it appears as if residential loads can be reasonably accurately modelled as constant current loads [34]. In this work, it is assumed that the loads on a residential feeder is constant current. The impact of this assumption can be the topic of future research.

2.3.3 Behaviour of loads with time: load profiles, load distributions and correlation of loads

External (weather) and internal driving forces (habits) influence the use of appliances. Individual consumers may respond differently to these forces, but a definite trend is visible when a large number (>30) of consumers is investigated. A load profile is a graphical representation of the trends in appliance use. Periods with similar trends are grouped together to form typical load profiles. The following issues specify residential load profiles:

- Type of day: Weekday, Saturday and Sunday
- Seasons: Winter, Summer, Spring/Autumn
- Load growth (see section 3.4)

Figure 6 shows a typical winter weekday load profile for a medium income, South African, community (Claremont) as measured in 1997. Note the following two curves on the graph:

- An **average** load curve, this is the average load per customer for the community during winter months
- An **upper limit** load curve, this should be interpreted as the maximum average load which can be expected for a specific half-hour.

The extreme of the upper limit curve is equal to the **ADMD** of the community. Claremont registered an ADMD of 4 kVA in 1997.

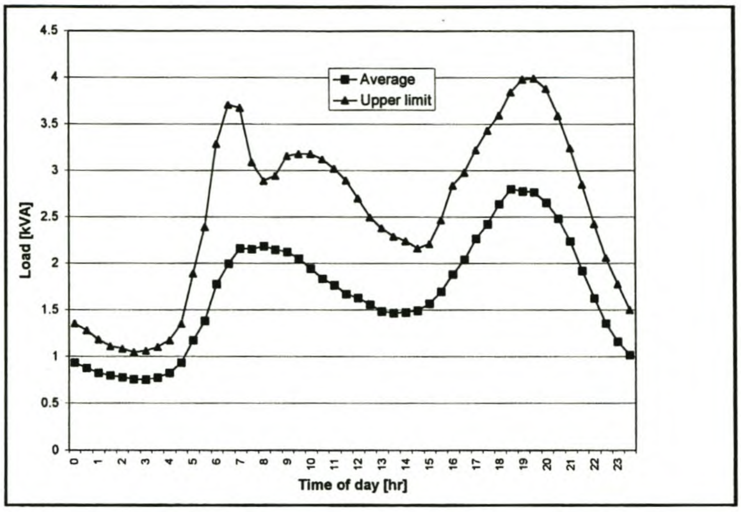

Figure 6 Typical winter weekday load profile for a Claremont as measured in 1997

The ADMD is the maximum average load of a number of consumers, each having a load corresponding to the appliances being used. Figure 7 shows a histogram of the individual loads and illustrates that the load currents at the instant of **system** maximum demand vary significantly.

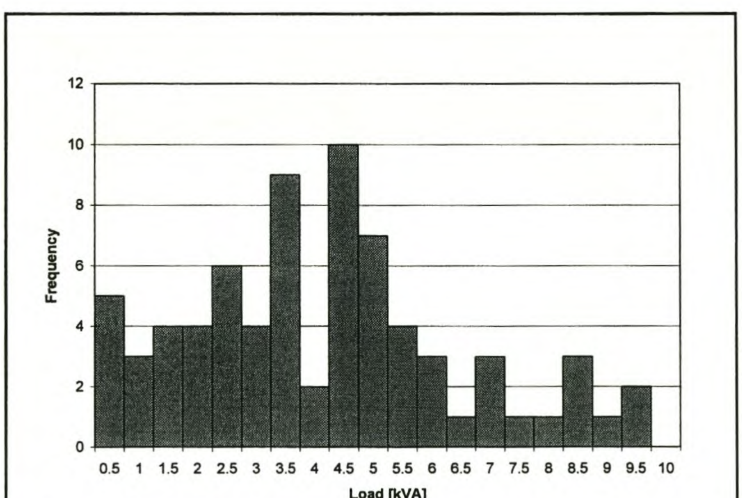

Figure 7 A histogram of the individual loads at the moment when the average load is maximum

The most significant statistical properties of this range of loads are the following:

- the mean, i.e. the average load in the range
- the standard deviation which is a measure of the spread of the distribution
- the minimum value, under normal situations it can be assumed to be zero
- the maximum value

These properties can be captured in one function by fitting a probability distribution function (pdf). The modelling of the load currents at the time of system maximum demand as a probability distribution and the manipulating of the moments of this distribution is the basis for the Herman Beta voltage calculation method [33]

The loads at any specific moment in time are caused by the appliance use of the individual consumers. The loads at any specific moment in time are related to the loads at any other moment in time through the specific appliance penetration of the different consumers. This relationship between loads at different instances can be measured in statistical terms as correlation or covariance. Figure 8 shows a scatter plot of the individual loads at two system peaks measured in Claremont. The correlation between the loads in the two events is estimated at 64%.

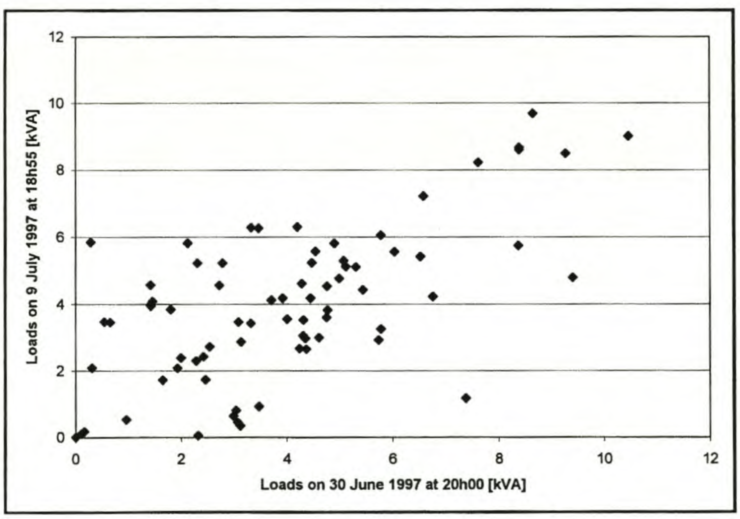

Figure 8 A scatter plot of the individual loads at two system peaks measured in Claremont 1997

Figure 9 shows three average profiles from three different consumers as measured in July 1997 in a medium income community (Claremont). Two timeslots with high load are marked on the graph. This is an oversimplified version of the previous scatter plot, but it should be clear why correlation between the different times of high load is present.

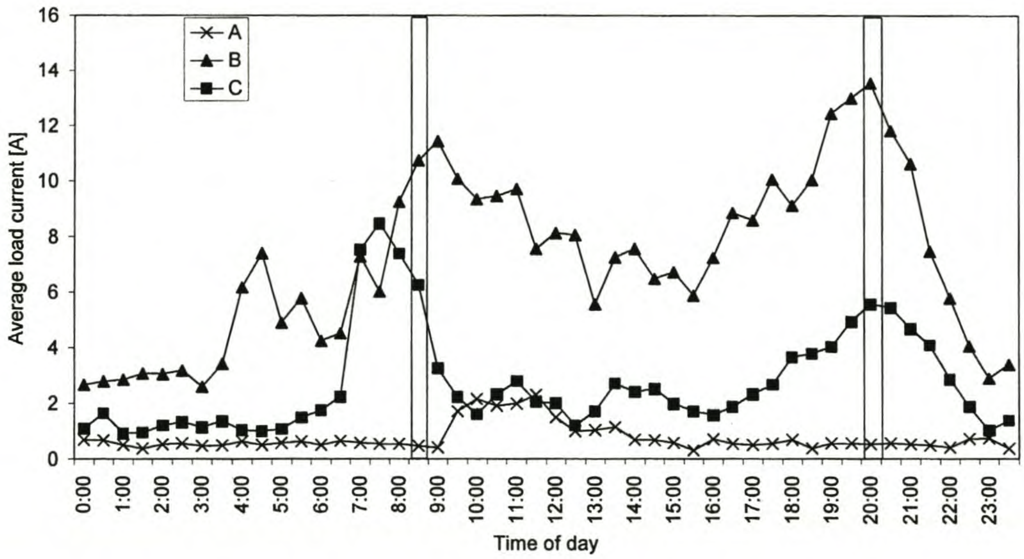

Figure 9 Three average load profiles of three different consumers as measured in Claremont in July 1997

Note that the loads at a specific moment in time are statistically **independent**, but the same loads considered over time are statistically **dependent**.

Figure 10 shows three histograms drawn from the above average load profiles. Instead of modelling the load currents at specific instants in time, each individual consumer's load can be described with a probability distribution. This probability distribution is equivalent to a load duration curve for a specific consumer, e.g. a 95% confidence level means the load is less than the percentile value for 95% of the time.

Each individual consumer load current therefore has a distribution with a mean, standard deviation, minimum and maximum value. Random variables of these distributions are correlated.

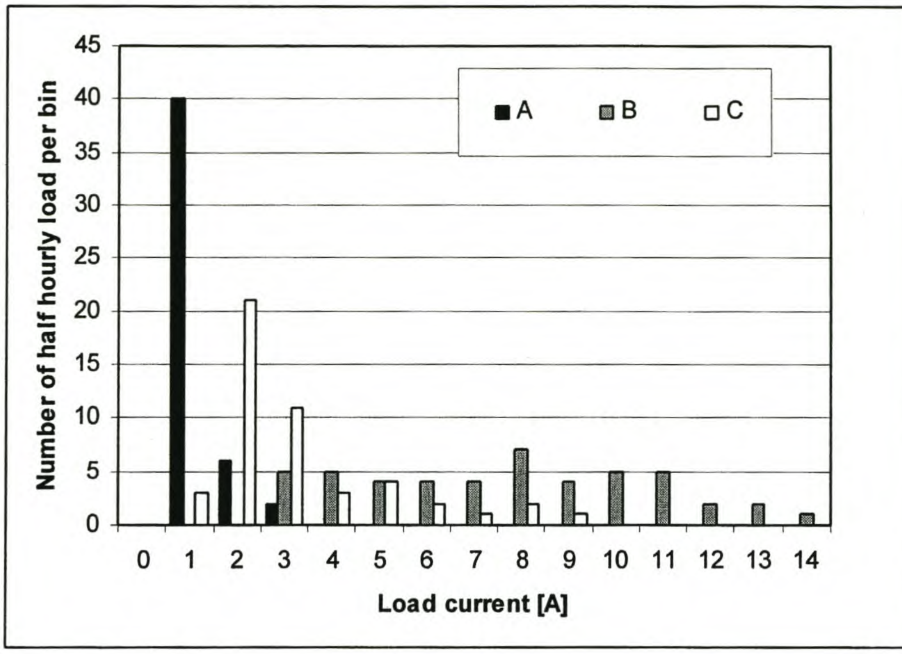

Figure 10 Histograms of the three average load profiles of three different consumers as measured in a medium income community (Claremont) in July 1997

If the correlation between the random variables is known, probabilistic load flow calculation can be made. Several probabilistic load flow calculation techniques can be found in literature [1,2,5,6, 39,57].

Unfortunately these methods cannot be used for residential low voltage feeder design. All of these methods assume that a distribution is known for each of the loads on the network. These distributions might be continuous or discreet but have to be specified before the load flow calculations can be performed. In low voltage feeder design very little is known about the consumers on the feeder. It is also very expensive to gather information about individual households.

The amount of information available at design time about each consumer is illustrated in the following figure. Aerial photographs of unserviced areas can give some indication of the size and shape of dwelling but very little more. Note that the three load profiles (A,B and C) and their associated distribution are known, but it is unknown whether the load current from house 1 is load A or load B, etc.

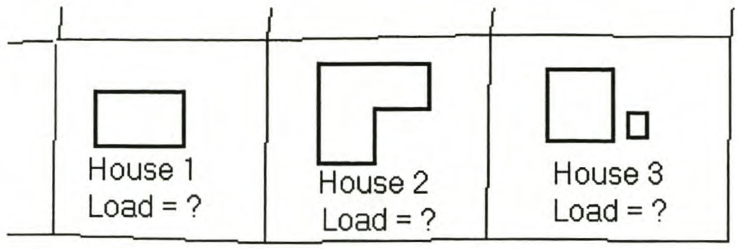

Figure 11 An illustration showing the type of information that is available at design time about individual residential consumers

If it assumed that each of loads A,B and C was measured at either house 1,2 or 3 then six combinations of house number and load number exists as illustrated in the following table:

Table 1 Possible assignments of loads A, B and C to houses 1, 2 and 3

Possible assignments	House number		
	1	2	3
Possibility 1	A	B	C
Possibility 2	A	C	B
Possibility 3	B	A	C
Possibility 4	B	C	A
Possibility 5	C	A	B
Possibility 6	C	B	A

If many loads and houses are present, a continuous range of probable load distributions might better represent the number of possible assignments of loads and houses. This implies that some information about the probability of finding a specific load distribution at a specific house is known. In the following chapters it will be shown that most of this information is contained in the distribution of consumption of individual consumers.

2.4 UNCERTAINTY IN MODELS AND PARAMETERS

2.4.1 Uncertainties in fitting statistical models

The proposed probabilistic models are based on some of the most significant statistical properties of a process or variable. Any model is subject to errors that are introduced when the model is fitted. Figure 12 shows a histogram of domestic loads and a fitted probabilistic model. The model is a probability distribution function (pdf) and specifically the beta pdf.

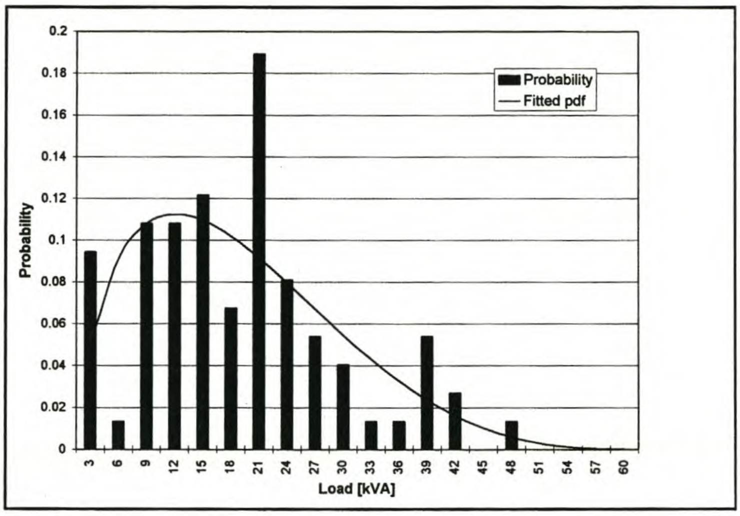

Figure 12 A histogram of domestic loads with a fitted probability distribution function

The difference between the fitted curve and the histogram is a measure of the **goodness of fit**. Two tests exist for estimating the goodness of fit [38]:

- Chi square test
- Kolmogorov-Smirnov test (KS test)

The KS test can be used to test any probabilistic model, since it literally measures the difference between the modelled probability and the actual probability. A method to deal with this type of uncertainty was suggested by [36] and uses fuzzy numbers to express the amount of belief or goodness of fit.

2.4.2 Load model parameter uncertainty

Load model parameters for design purposes are derived through one of the following two processes:

- prediction
- forecasting

Prediction: a model of the relationship between one or more attributes (predictors) and electrical parameters are used to predict design parameters. The following can cause errors in the prediction:

- An error in estimating a particular *predictor*, e.g. errors in estimating the household income of a community, produce incorrect estimates of the ADMD when using the South African Pre-electrification Tool 1998 [17].
- A poor or ill-conditioned fit between the predictors and the prediction contributes to the total error of the prediction.

Forecasting: measurements from historical load data are used to predict future load parameters. The following contributes to the total forecast uncertainty:

- Accuracy of the measurements
- Applicability of load data
- The models used to relate the past with the future

Either method of specifying the design parameters are uncertain. Fuzzy numbers can be used to account for this type of uncertainty and it was used by [63] to specifically deal with the specification of design parameters.

2.4.3 Network parameter uncertainty

Un certainties about the network parameters specified at design time are the following:

- *The cable resistance due to the uncertainty of the exact length of the cable*: the cable length can vary due to sag, differences between planned and actual layouts, etc.

- *The cable resistance due to the uncertainty in the temperature of the conductor*: the conductor resistance is influenced by the ambient temperature and the load current (loss).
- *The supply voltage to the feeder:* the voltage supplied to a LV feeder can vary due to (a) voltage regulation in the MV network and (b) voltage regulation over the MV/LV transformer.

All of the above can be specified in terms of a minimum, maximum and most likely value, which is ideal for modelling as fuzzy number. This approach has been used by [19].

The temperature rise in cables due to load current, ambient temperature and other external factors is not part of this study. The change in resistance of the cable due to the temperature can be accurately modelled if the proper thermal equations are applied. This could be the topic of further research.

2.5 PROPOSED UNCERTAINTY MODEL

Load uncertainties are dealt with using probabilistic methods, but the load parameter uncertainty and the network parameter uncertainties are best modelled using fuzzy numbers. An uncertainty model with a fuzzy and probabilistic component is therefore required to express the total uncertainty.

A fuzzy probability distribution function is described in [20] and a range of possible values is present at each level of confidence. This approach is proposed to represent the total uncertainty due to

- Probabilistically modelled loads
- Fuzzy load and network parameters

An example of the application of this approach can be found in [63]. Figure 13 shows a fuzzy cumulative probability distribution function of load current.

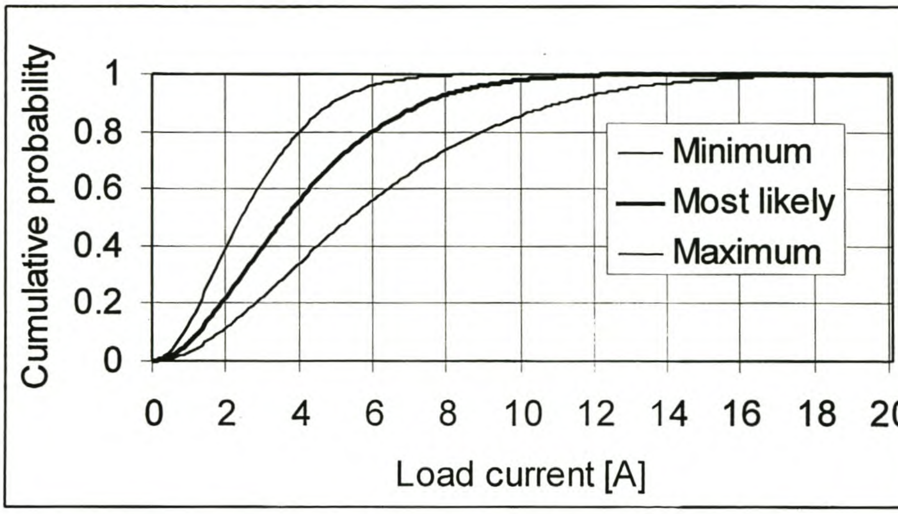

Figure 13 A fuzzy cumulative probability distribution of load current

At a 90% confidence level, the percentile value is a fuzzy number with

- Minimum = 4.9 A
- Most likely value = 7.4 A
- Maximum = 11.1 A

The fuzzy range of values can be changed to a crisp range by applying and α-level cut to the fuzzy percentile value.

This model is the basis to the method described in chapter 6.

2.6 SUMMARY

This chapter gives an overview of the methods that can be used to quantify uncertainty. The uncertainties considered in this document are residential load uncertainty, load parameter and network parameter uncertainty.

Residential load uncertainty can be modelled with probabilistic methods and these methods will be used to manipulate the load uncertainty into voltage performance and loss uncertainties.

The certainty about the calculated of voltage performance and loss is reduced by network and load parameter uncertainty. These uncertainties can best be modelled using fuzzy techniques.

A combined fuzzy and probabilistic uncertainty model is therefore proposed to represent the total load uncertainty.

3. RESIDENTIAL CONSUMER LOAD MODEL

3.1 INTRODUCTION

A probabilistic load model for residential consumers is introduced in this chapter. This model forms the basis for the probabilistic methods introduced in chapters 5 and 6. The model represents the load currents for a community over a period of time. The load current of each individual consumer can be represented with an average value and a standard deviation value. The load currents of individual consumers are correlated due to the following common influences:

- Daily living habits
- Weather
- Cultural traditions

The load model further represents the uncertainty of the load parameters for each individual consumer.

3.2 LOAD MODEL

Figure 14 shows the average load current trace for two consecutive days as measured in Tambo during a TSI load research project in 1998 [61]. The trace has a mean value (indicated by the bold horizontal line) and some movement away from the mean. This movement can be measured as the standard deviation of the trace. Two thin horizontal lines show the 10^{th} percentile and 90^{th} percentile of the trace.

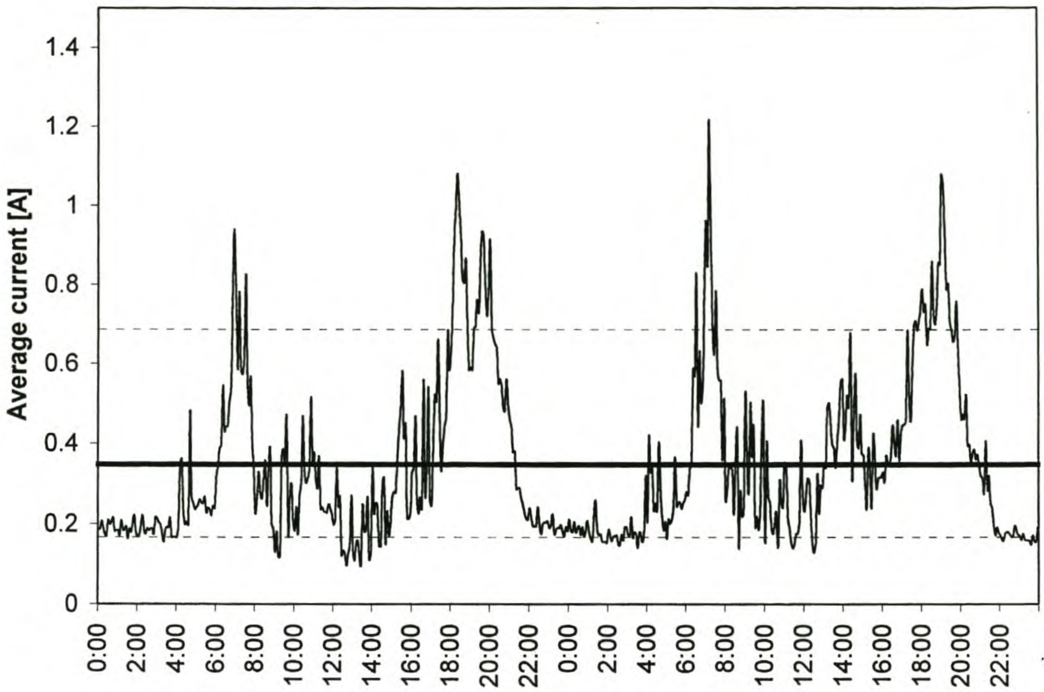

Figure 14 The average load profile as measured in two consecutive days in Tambo during 1998

Over a longer period of time, the pattern is repeated, with some days higher and others lower. The variation is caused by residential consumer's habits, time of day, climate etc. This causes the loads to be dependent and correlation needs to be taken into account when manipulating the load model.

Each consumer's load current for a period of time has a distribution, which can be described with a mean, μ_k and a standard deviation, σ_k. Individual consumers' loads are correlated, and the correlation between consumer k and consumer m is ρ_{km}.

The shape of the distribution of the individual consumers is non-Gaussian and in most cases is left skewed. Figure 15 shows a histogram that was calculated from an average current trace as measured in Tambo. The histogram shows that the distribution is left skewed and a beta distribution was fitted to the histogram (shown on the graph as a solid line). The goodness of fit of the distribution was rejected at α=0.01 using the Kolmogorov-Smirnov test for goodness of fit [38].

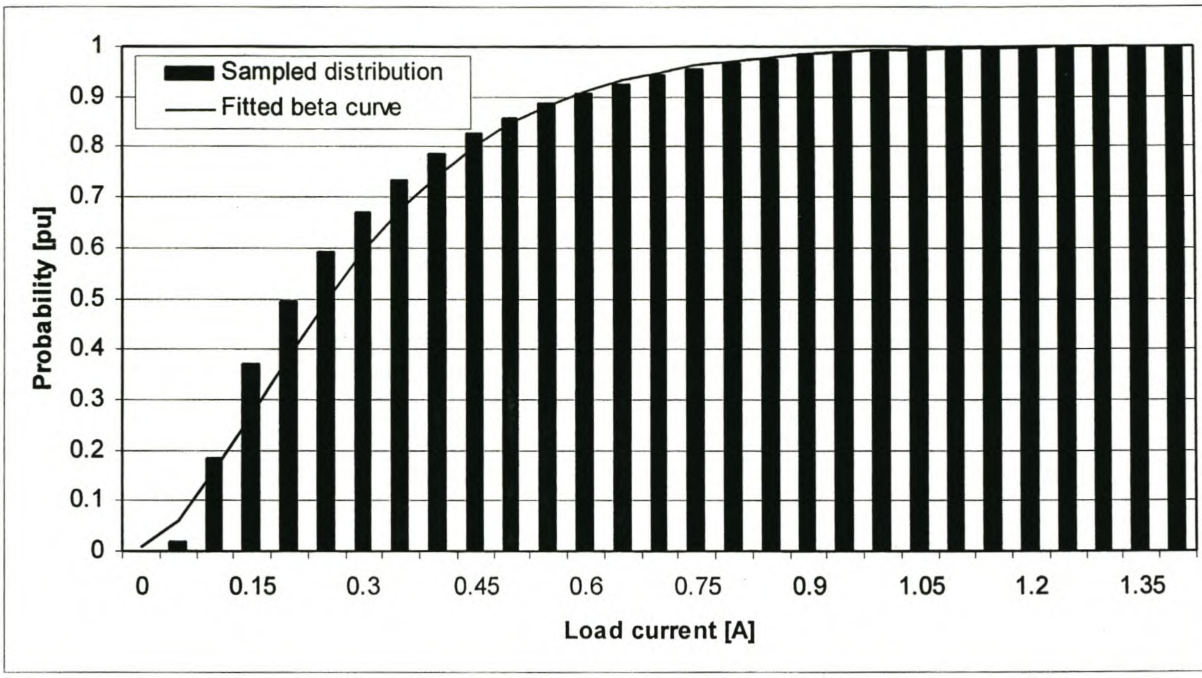

Figure 15 Histogram of the average load current in Tambo

The curve is however useful for estimating percentile values at confidence values greater than 90%. Thus, assuming a beta probability distribution, the percentile value at some level of confidence is:

$$X_{\%tile} = \text{qbeta}(Cl,\alpha,\beta).(H-G)+G+e_x \tag{3}$$

where

$X_{\%tile}$	is a percentile value at confidence level cl
Cl	is the confidence level (pu)
Qbeta	is the inverse beta function
H	is the maximum value of X
G	is the minimum value of X
α,β	is the parameters for the beta distribution function
e_x	is the error between the actual percentile value and the real percentile. The average of e_x is zero and has a variance $V(e_x)$.

For some community of residential consumers, a range of possible load parameters might be known. The effect of e_x can be ignored if $V(e_x)$ is small and is randomly distributed between the consumers in the community, i.e. for some consumers the percentile is over- estimated and for some other consumers the percentile is under-estimated. This was found to be the case for medium and low income consumers in South Africa. If the sum of a number of consumer loads is being estimated, the effect of e_x decreases.

Proposed load model:

A range of possible load parameters might be known for a community, but the load parameter of a specific consumer is uncertain. Effectively, the methods presented in the following sections calculate the mean and variance of the effects due to a range of load parameters (see below). To derive the methods, models for individual loads are required and these are assumed to be:

- Each consumer's load trace has a specific mean value, μ_k, and standard deviation, σ_k
- Tail-end percentile values can be estimated using a beta distribution
- Between any two load traces, correlation exists and is expressed through Pearson's correlation coefficient

Similar load models have been used in load flow studies [1, 2 455-507, 5, 6, 39,57]. The use of the beta distribution in this type of study is not very common. A probabilistic voltage regulation method for residential consumers using the beta distribution was proposed by [33] following a comparison of the appropriateness of a range of probability distribution functions by [31]. The beta probability distribution function has a number of properties which make it ideal for the modelling of South African residential load current distributions.:

- It can be left-skewed (typical of consumers whose load are not restricted)
- It can be bath-tub shaped or right-skewed (typical distributions for load restricted consumers 2.5 A to 10 A circuit breaker)
- It has a finite base (corresponding to 0 and full circuit breaker load)

Residential loads can be approximately modelled as constant current [34] in static load flow studies. The proposed load model assumes that the loads are constant current for all consumer groups. For very low income consumers, this assumption might give over-estimates of the voltage regulation and losses. At the time of writing no load data were available to test this assumption or calibrate the proposed procedures. Future research should be performed to verify this assumption.

3.3 PROPERTIES OF THE LOAD MODEL

This section gives a brief overview of the properties of random variables that were used to derive the methods found in later chapters. The properties of the sum of two consumers can be calculated using the following equations:

- Mean for two consumers

$$\mu_{total} = \mu_1 + \mu_2 \tag{4}$$

- Standard deviation for two consumers

$$\sigma_{total} = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2} \tag{5}$$

ρ_{12}= correlation between consumer 1 and consumer 2

Similarly, for N consumers:

- Mean for N consumers

$$\mu_{total} = \sum_{i=1}^{N} \mu_i \tag{6}$$

- Standard deviation for N consumers

$$\sigma_{total} = \sqrt{\sum_{i=1}^{N} \sigma_i^{\,2} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \rho_{ij} * \sigma_i * \sigma_j} \tag{7}$$

ρ_{ij}= correlation between consumer i and consumer j

These relationships form the basis of the methods derived in section 4 and 5.

3.4 DISTRIBUTION OF LOAD MODEL PARAMETERS

3.4.1 Individual consumer load model parameters

Figure 16 shows a histogram of the average current for individual consumers as measured in Tambo township in 1998.

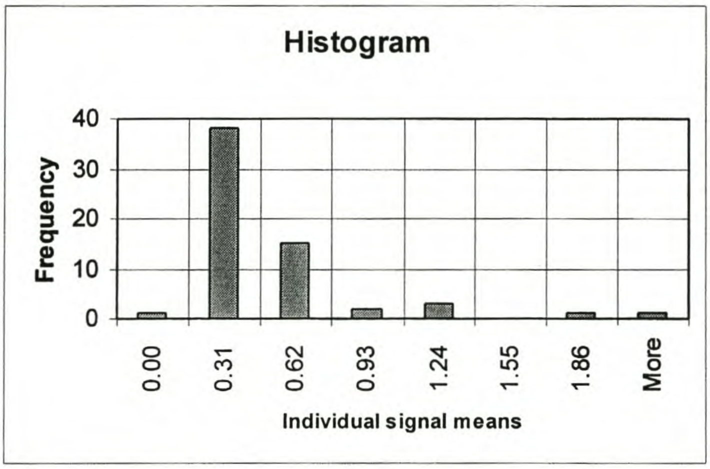

Figure 16 Histogram of the average current for individual consumers taken from a low income community (Tambo)

In each community a range of mean values can be measured. The more varied the consumers are (in terms of habit and appliance ownership) the wider the standard deviation of this distribution. This distribution can be specified with a mean and a standard deviation. Section 3.9 contains a table with typical values for different communities. The table gives some indication of the income (very low, low, medium and high) for the different communities.

The standard deviation of the individual consumers also has a distribution, which is correlated with the individual mean distribution as can be seen in the following scatter plot. The scatter plot shows the relationship between the average current per consumer and the variance of the associated consumer's current trace.

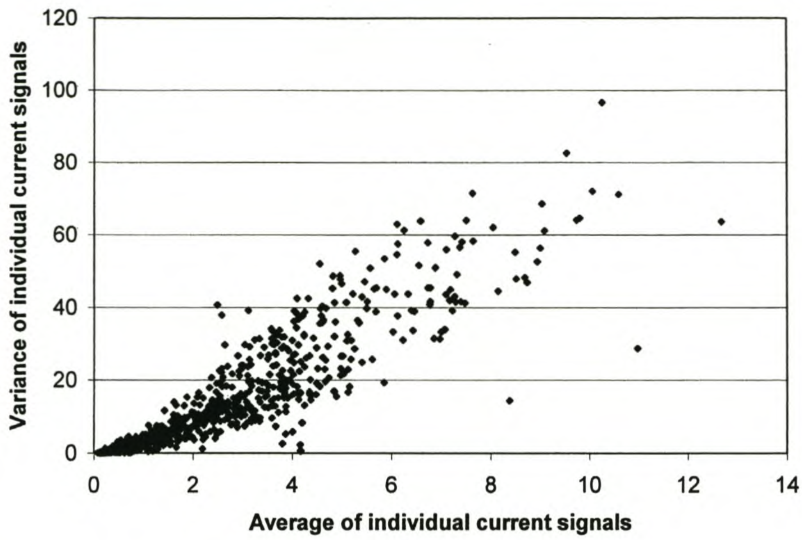

Figure 17 A scatter plot showing the relationship between the individual current means and variances for 823 consumers from various income groups as stored in the NRS LR database [53]

The relationship between the standard deviations and the averages of the consumer load traces can be expressed as

$$\sigma_k = G.(\mu_k - \mu_{total}) + C_o + e_k \qquad (8)$$

where:

G	is the slope of a linear regression
C_o	is the intercept of a linear regression
e_k	is an error term with zero mean and $V(e_k)$=(standard error of the regression)^2
σ_k	is the standard deviation of the current for consumer k
μ_k	is the mean of the current for consumer k
μ_{total}	is the mean of the individual currents of a group of consumers to which consumer k belongs

Similarly, the linear component of the relationship between the variances and the averages of the consumer load traces can be expressed as:

$$\sigma^2_k = G_2.(\mu_k - \mu_{total}) + C_{o2} + e_k \qquad (9)$$

where:

G_2	is the slope of a linear regression
Co_2	is the intercept of a linear regression
e_k	is an error term with zero mean and $V(e_k)$=(standard error of the regression)^2
σ^2_k	is the variance of the current for consumer k
μ_k	is the mean of the current for consumer k
μ_{total}	is the mean of the individual currents of a group of consumers to which consumer k belongs

Note that the difference between the individual mean and the group mean is used as the independent variable in the regressions.

In typical applications of the load model, the load parameters of consumers at various points on the network are known, through measurements or some assumption about load class. In this application the exact load parameters of consumers connected to a feeder is unknown because very little is known about the consumers. Some work has been done to evaluate the influence of various socio-demographic factors on individual load behaviour and to date the strongest link found was through household floor area and geyser[16]. However, none of these models are very general and were applied only at a group level.

The parameters of the load model for each consumer can therefore assume a range of possible model parameters, for which a probability distribution might be known. If the load parameters probability distribution is known for individual consumers are known, the method in section 4, may be used to evaluated the voltage regulation. On the other hand, if the distribution is unknown or cannot be estimated, a fuzzy probabilistic method as described in section 6 should be used.

3.4.2 Group load current parameters

The after diversity demand load trace of a group of consumers also has a mean (mu) and a variance (σ_n). Here, σ_n is related to the standard deviation and variance of individual consumers through the following approximation:

$$E[o_n{}^2] \approx \frac{N * E[o_1{}^2] + N(N-1)E[o_1]^2 * E(\rho)}{N^2} \qquad (10)$$

where

$E(\sigma^2{}_n)$ is the expected value of the variance of the average current trace of N consumers

$E(\sigma^2{}_1)$ is the expected value of the individual variances $E(\sigma_1)$ is the expected value of the individual standard deviations

$E(\sigma_1{}^2)$ is the expected value of the individual variances

$E(\rho)$ is an average correlation term

Note: It is assumed that ρ and the individual standard deviations are uncorrelated or usefully uncorrelated. This was tested for low income and medium income consumers and could cause an over-estimate of 5%. The test was in the form of a Monte Carlo type simulation. The over-estimate was not found in the simulations when data from a low-income community was used.

Equation 10 may be used to estimate the expected value of ρ, given some measurement of the group load behaviour when N is sufficiently large.

3.5 MAXIMUM CURRENT

3.5.1 Maximums as percentile values

The maximum value that a load current trace will assume may be estimated using the consumer's load current probability distribution function. The maximum value of the current is the percentile value of the distribution at some level of confidence.

Given a current trace with some mean μ_k and standard deviation σ_k, the percentile value at some level of confidence can be calculated using the following equations (assuming a beta probability distribution).

$$\alpha = \frac{\mu_{ns}{}^2 - \mu_{ns}(\sigma_{ns}^2 + \mu_{ns}^2)}{\sigma_{ns}^2} \qquad (11)$$

$$\beta = \frac{\alpha(1-\mu_{ns})}{\mu_{ns}} \tag{12}$$

where

$$\mu_{ns} = \frac{\mu_k - \text{Minimum}}{\text{Maximum} - \text{Minimum}} \tag{13}$$

$$\sigma_{ns} = \frac{\sigma_k}{\text{Maximum} - \text{mininimum}} \tag{14}$$

In the final step of this calculation, $I_{\%tile}$ is obtained by scaling the result of the inverse beta probability function with the minimum and maximum values.

$I_{\%tile}$= qbeta (Confidence value,α, β)*(Maximum-minimum) +minimum

where

μ_k	is the mean value of load current trace k
σ_k	is the standard deviation of load current trace k
α_k, β_k	are the parameters of the beta pdf
μ_{ns}	is μ_k scaled to the range 0 to 1
σ_{ns}	is σ_{ns} scaled to the range 0 to 1
Maximum	is the upper bound of the scaling range
Minimum	is the lower bound of the scaling range
Qbeta	is the inverse beta function

Note: the circuit breaker size is a convenient value for the maximum of the scaling range. For residential consumers, the minimum value is typically 0. This means that the *maximum* is calculated as

$$\text{Maximum} = N.C_k \tag{15}$$

Where N is the number of consumers and C_k is the circuit breaker size

3.5.2 Distribution of percentile values of individual consumers

The percentile value (at some level of confidence) varies between consumers in a community and when measured, would take on a range of values. The range of percentile values can be described by a probability distribution function with a mean and a standard deviation. The variation in the percentile value is not the same as an extreme value distribution but reflects the differences between consumers connected to the same feeder.

In order to calculate the distribution of percentile values, the following information is required about a group of consumers:

- The average current for each consumer, μ_k
- The standard deviation of the current for a consumer, σ_k
- The correlation between different consumers or the standard deviation of the average current trace for a group of N consumers, where N is large (N>100). The correlation coefficient can then be estimated using (6).

The average value of the percentile distribution for one consumer can be calculated as follows:

i. Calculate the expected value of the consumer current means, this is $\mu_{average}=E(\mu_k)$

ii. Calculate the expected value of the consumer current variance, this is $E(\sigma_k^2)$

iii. Calculate the minimum and maximum values and use this to obtain μ_{ns} and σ_{ns}, by scaling $\mu_{average}$ and $E(\sigma_k^2)$

$$\mu_{ns} = \frac{\mu_k - \text{Minimum}}{\text{Maximum} - \text{Minimum}} \tag{16}$$

$$\sigma_{ns} = \frac{\sigma_k}{\text{Maximum} - \text{mininimum}} \tag{17}$$

Note that Minimum = 0 and Maximum = $N.C_k$ with C_k the circuit breaker size

iv. Calculate the α, β using μ_{ns} and σ_{ns}

$$\alpha = \frac{{\mu_{ns}}^2 - \mu_{ns}(\sigma_{ns}^2 + \mu_{ns}^2)}{\sigma_{ns}^2} \tag{18}$$

$$\beta = \frac{\alpha(1 - \mu_{ns})}{\mu_{ns}} \tag{19}$$

v. Use the inverse of the beta distribution function, qbeta to approximate the percentile value, $I_{\%\text{-tile}}$

Note that this procedure does not give the expected value of the percentile values, but is an approximation. The approximation was tested on 823 consumers from the NRS LR database and the results are within 0.1 A of the actual average of the percentile values.

The spread of the percentile distribution for one consumer can be calculated as follows:

vi. Calculate the variance of the consumer current means, $V\delta$, from the source data. Note that $V\delta$ is also equal to $E(\delta_k^2)$ and

$$\delta_k = \mu_k - E(\mu_k) \tag{20}$$

Where

μ_k is the average of the current trace for consumer k

$E(\mu_k)$ is the expected value of the average current traces for a group of consumers

δ_k is the difference between μ_k and $E(\mu_k)$

vii. Calculate values for G, G_2, C_o and C_{o2}, the slopes and intercepts of the relationships between the variances, standard deviations and δ_k's of the individual load currents, using linear regression:

$$\sigma^2{}_k = G_2 * \delta_k + C_{o2} + e_{k2} \tag{21}$$

$$\sigma_k = G * \delta_k + C_o + e_k \tag{22}$$

See section 7 for more details on obtaining G_2, C_{o2}, G and C_o. Figure 18 shows an example of a regression line fitted with data measured in a medium income community in South Africa (Claremont). The regression explains ±70% of the variation in the source data.

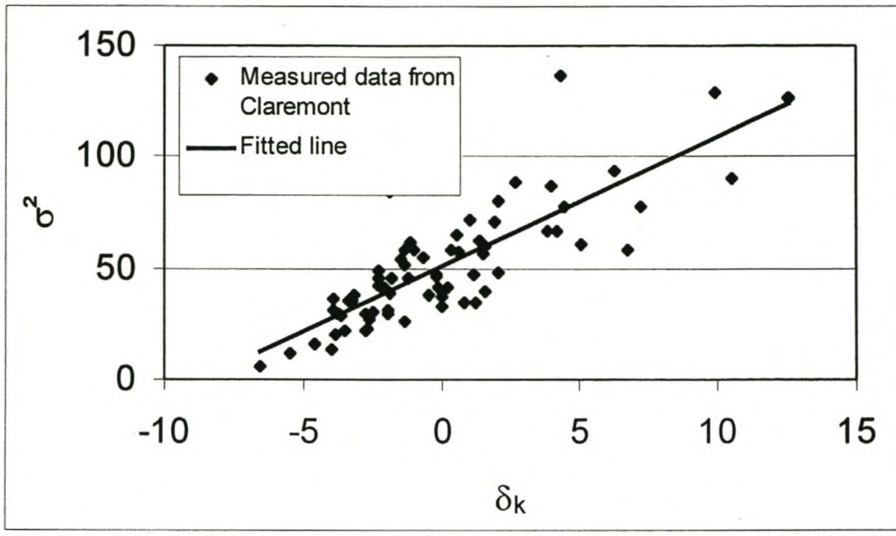

Figure 18 **Fitted line between δ_k and σ_k^2 for data as measured in a medium income community (Claremont)**

viii. The variance of the percentile distribution due to the variance in δ_k is given by the following equation:

$$V(I_{\%tile} \mid \delta) = \left[\frac{\partial}{\partial \delta_k} qbeta(cl, \alpha(e_k), \beta(e_k)) * C_k \right]^2 V\delta \tag{23}$$

where

$V\delta$ is the variance of δ_k

$V(I_{\%tile}|e_o)$ is the variance of the percentile distribution due to the variance in δ_k, i.e. the marginal variance of $I_{\%tile}$ given e_o

ix. The variance of the percentile distribution due to the error in the regression is given by the following equation:

$$V(I_{\%tile} \mid \delta) = \left[\frac{\partial}{\partial e_k} qbeta(cl, \alpha(e_k), \beta(e_k)) * C_k \right]^2 Ve_o \tag{24}$$

where

Ve_o is the square of the standard error of the linear regression

$V(I_{\%tile}|\delta_k)$ is the variance of the percentile distribution due to the variance in e_k, i.e. the marginal variance of $I_{\%tile}$ given δ

x. The total variance is the sum of (vii) and (viii) since δ_k and e_k is independent.

Steps (vii) and (viii) can be evaluated using numerical methods and are based on a linearization through a first order Taylor series [2 p 130-131, 39]. The probability distribution is obtained by fitting a beta probability distribution to the mean (iv) and the variance, (ix) as described above.

3.5.3 Distribution of percentile values of a linear combination of consumers

A linear combination of consumers can be expressed as:

$$\sum_{i=1}^{N} a_i I_i \tag{25}$$

where a_i is a coefficient associated with current trace i.

The distribution of percentile values of a linear combination of consumers can be calculated using the following steps:

i. First calculate the expected value of the linear combination of load currents:

$$E\left(E\left(\sum_{i=1}^{N} a_i I_i\right)\right) = \mu \sum_{i=1}^{N} a_i \tag{26}$$

ii. Calculate the expected value of the variance of the linear combination:

$$E\left(V\left(\sum_{i=1}^{N} a_i I_i\right)\right) = \sum_{i=1}^{N} a_i E(o^2) + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} a_i a_j E(o)^2 E(\rho) \tag{27}$$

where

$E(\sigma^2)$ is the average of the variances of the individual current traces

$E(\sigma)$ is the average of the standard deviations of the individual traces

$E(\rho)$ is the average of the correlations between individual current traces.

This assumes that the correlations and variances are independent.

Note that the expected value of standard deviation of the linear combination is required to approximate the expected value of the percentile values using equation (7) and (8). This means that this method over-estimates the expected value of the percentile values. The extent of the over- estimation is dependent on the number of consumers (the more consumers, the smaller the over-estimate) and the differences between the consumers in the group. None of the current test cases show problems due to this over-estimate.

iii. Calculate the minimum and maximum values of the linear combination:

$$\text{Minimum} = \sum_{a_i<0} a_i C_k \tag{28}$$

$$\text{Maximum} = \sum_{a_i>0} a_i C_k \tag{29}$$

where

C_k is the circuit breaker size or maximum value of the load current for consumer k

iv. Calculate the expected value of the linear combination of percentile values

Using the equations as detailed in section 3.5.2, calculate the expected value of the percentile values.

v. Calculate the variance of the linear combination of percentile values

The variance of the linear combination of percentile values due to the variance between the load trace means is obtained using the following sum of partial derivatives of the equations used in steps i-iii.

$$V(I_{\%-tile} \mid e_o) = \sum_{i=1}^{N} \left[\frac{\partial}{\partial \delta_i} qbeta(cl, \alpha(\delta_i), \beta(\delta_i)) * C_k \right]^2 V\delta \qquad (30)$$

The variance of the linear combination of percentile values due to the errors in the regression is obtained using the following sum of partial derivatives of the equations used in steps i-iii.

$$V(I_{\%tile} \mid \delta) = \sum_{i=1}^{N} \left[\frac{\partial}{\partial e_i} qbeta(cl, \alpha(e_i), \beta(e_i)) * C_k \right]^2 Ve_o \qquad (31)$$

Section 3.8 contains an implementation of these equations on a Mathcad 2000 worksheet.

3.6 TESTS

The methods described in the previous sections were tested using Monte Carlo type simulations. The purpose of the test is to compare simulated 90th percentiles and calculated values. The procedure for the test were as follows:

1. Select from the measured loads in a community, N consumer load traces.
2. Add the load traces, for each period of measurement (5 minutes)
3. Calculate the percentile value of the combined load trace at 90% confidence
4. Repeat steps 1 to 3, 500 times
5. Calculate the average and standard deviation of the 500 percentile values
6. Repeat steps 1 to 5, for N=1,10 and 15

The Monte Carlo simulation was performed using load data from Claremont for the month of July and using load data from Tambo for five months (April – August, inclusive).

The following table shows a comparison of the calculated and simulated results for Claremont, a community with medium income.

Table 2 The measured and calculated current 90th percentiles for Claremont and Tambo

Number of consumers	Monte Carlo			Calculated		Difference	
	Average [A]	Standard deviation [A]	Percentage error at 90% confidence	Average [A]	Standard deviation [A]	Average [A]	Standard deviation [A]
Claremont							
1	16.5	6.1	2.7	17.4	6.3	- 0.8	- 0.2
10	119.1	15.8	1.0	120.5	15.0	- 1.4	0.7
15	174.7	20.3	0.9	176.4	18.1	- 1.7	2.2
Tambo							
1	0.6	0.9	10.0	0.9	1.3	- 0.3	- 0.5
10	7.4	2.8	2.8	7.6	2.6	- 0.2	0.2
15	11.1	3.2	2.1	10.6	2.9	0.4	0.2

The maximum percentage error obtained was 10% at 90% confidence.

The differences between the calculated and Monte Carlo results can be attributed to (see section 3.10 for a general discussion of results):

- The small number of Monte Carlo iterations (500)
- The small number of consumers from each community (60-70)

One of the assumptions in the calculation method is that the distribution of consumers is continuous. However, in the Monte Carlo simulation only between 60 and 70 consumers from this distribution were available, resulting in a discreet distribution. This could explain the difference between the simulated and calculated results for Tambo where one consumer is considered.

3.7 SUMMARY AND CONCLUSION

A probabilistic model for residential consumer loads is described in this chapter. The model uses the beta probability distribution to represent the load currents of individual consumers over a period of time.

A method for calculating the maximum value of a linear combination of load currents was described. This method estimates the expected value and variance of a percentile of the load current distribution and uses the percentile value to approximate the maximum load current.

3.8 APPENDIX: MAXIMUM CURRENT SAMPLE CALCULATIONS

This section contains an implementation in MathCAD 2000 of the method described in section 3.5.3. A set of parameters and results is given for the various variables used.

$N = 10 \qquad c_k = 20 \qquad \rho = 0.07$

$V\delta = 0.15 \qquad \mu = 0.322 \qquad cl = 0.9$

$G = 1.41 \qquad C_o = 0.63 \qquad Ve_o = 0.366^2$

$G2 = 3.89 \qquad C_{o2} = 0.82 \qquad Ve_{o2} = 0.336^2$

$\delta = 0 \qquad e_o = 0$

$$a = \begin{pmatrix} 1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1\\1 \end{pmatrix}$$

$\delta_k(\delta,k,i) = \text{if}(k = i,\delta,0) \qquad e_k(e_o,k,i) = \text{if}(k = i,e_o,0)$

$\mu_k(\delta,k,i) = \mu + \delta_k(\delta,k,i)$

$\sigma_k(\delta,e_o,k,i) = \left(G\cdot\delta_k(\delta,k,i) + C_o + e_k(e_o,k,i)\right)$

$\sigma_{k2}(\delta,e_o,k,i) = \left(G2\cdot\delta_k(\delta,k,i) + C_{o2} + e_k(e_o,k,i)\right)$

$$\mu_n(\delta,k) = \sum_{i=0}^{N-1}\left(a_i\cdot\mu_k(\delta,k,i)\right)$$

$$\sigma_n(\delta,e_o,k) = \sqrt{\sum_{i=0}^{N-1}(a_i)^2\left(\sigma_{k2}(\delta,e_o,k,i)\right) + \sum_{i=0}^{(N-1)}\sum_{j=0}^{(N-1)}(i \neq j)\cdot\left(a_i\cdot a_j\cdot\sigma_k(\delta,e_o,k,i)\cdot\sigma_k(\delta,e_o,k,j)\right)\cdot\rho}$$

$\text{MinimumCalc}(i,j) = \text{if}(a_{i,j} < 0, a_{i,j}, 0)$

$\text{MaximumCalc}(i,j) = \text{if}(a_{i,j} > 0, a_{i,j}, 0)$

$\text{Minimum} = \sum \text{matrix}(N,1,\text{MinimumCalc})\cdot c_k \qquad \text{Minimum} = 0 \qquad \mu_n(0,0) = 3.22$

$\text{Maximum} = \sum \text{matrix}(N,1,\text{MaximumCalc})\cdot c_k \qquad \text{Maximum} = 200 \qquad \sigma_n(0,0,0)^2 = 10.7$

$$\mu_{ns}(\delta,k) = \frac{\left(\mu_n(\delta,k) - \text{Minimum}\right)}{\text{Maximum} - \text{Minimum}} \qquad \mu_{ns}(0,0) = 0.016$$

$$\sigma_{ns}(\delta,e_o,k) = \left[\frac{\sigma_n(\delta,e_o,k)}{(\text{Maximum} - \text{Minimum})}\right] \qquad \sigma_{ns}(0,0,0)^2 = 2.675\times 10^{-4}$$

$$\alpha(\delta, e_o, k) = \frac{\mu_{ns}(\delta,k)^2 - \left(\sigma_{ns}(\delta,e_o,k)^2 + \mu_{ns}(\delta,k)^2\right) \cdot \mu_{ns}(\delta,k)}{\sigma_{ns}(\delta,e_o,k)^2} \qquad \alpha(0,0,0) = 0.937$$

$$\beta(\delta, e_o, k) = \frac{\alpha(\delta,e_o,k)\left(1 - \mu_{ns}(\delta,k)\right)}{\mu_{ns}(\delta,k)} \qquad \beta(0,0,0) = 57.278$$

$$\mathrm{Max}_k(\delta, e_o, k) = \mathrm{qbeta}\left(\mathrm{cl}, \alpha(\delta,e_o,k), \beta(\delta,e_o,k)\right) \cdot (\mathrm{Maximum} - \mathrm{Minimum}) + \mathrm{Minimum}$$

$$\mathrm{Max}_k(0,0,0) = 7.516011$$

$$\sum_{i=0}^{N-1} \left[\left(\frac{d}{de_o}\mathrm{Max}_k(\delta,e_o,i)\right)^2\right] \cdot Ve_o = 0.137 \qquad \sum_{i=0}^{N-1} \left[\left(\frac{d}{d\delta}\mathrm{Max}_k(\delta,e_o,i)\right)^2\right] \cdot V\delta = 6.344$$

3.9 APPENDIX: TYPICAL MONTHLY CONSUMPTION AND STANDARD DEVIATION VALUES

The following table contains typical values for the average consumption and standard deviation of consumption for different villages and townships in South Africa. An indication is provided of the income level in each of the communities (Very low, Low, Medium or High). All values are in kWh per month.

Name of village of township	Income	Average of consumption	Standard deviation of consumption per individual
Claremont 96	Medium	1,014.5	467.4
Claremont 97	Medium	990.7	407.5
Claremont 98	Medium	984.2	645.9
Cloetesville 94/5	Medium	567.6	172.1
Helderberg 97	High	844.1	302.6
Helderberg 98	High	782.0	295.6
Helderberg 99	High	853.3	319.5
Kwazakhele 95/6	Low	210.2	146.4
Lotus Park 99	Medium	684.6	182.3
Lotus Prk 98	Medium	719.7	252.3
Manyasteng 97	Very low	236.0	254.6
Manyatseng 96	Very low	238.7	341.6
Orient Hills 98	Low	517.9	201.8
Orient Hills 99	Low	494.9	172.4
Rontree Estate 99	High	1,077.2	423.1
Sanctuary Gdns 99	High	563.1	195.0
Summerstrand 99	High	657.5	372.7
Sweetwaters 96	Low	163.0	165.7
Sweetwaters 97	Low	267.9	314.5
Tafelsig 98	Low	424.9	204.9
Tafelsig 99	Medium	421.4	199.4
Umgaga 98	Low	238.1	195.6
Umlazi 95/6	Medium	176.4	114.1
Umlazi 98	Medium	490.2	285.5
Umlazi 99	Medium	477.4	280.9
Walmer 98	Very low	166.4	124.3
Walmer Est 97	Very low	146.5	128.3

3.10 APPENDIX : DISCUSSION OF THE MONTE CARLO SIMULATIONS USED FOR TESTING

3.10.1 Uncertainty in Monte Carlo results

Monte Carlo type simulations are used in several sections throughout this document to assess the results of the proposed calculation methods. The result of a Monte Carlo type simulation is normally specified as an average value and a standard error or absolute error and level of confidence, e.g. 5% error at 95% confidence.

Uncertainty about the results of any Monte Carlo type process decreases as the number of iterations increases, e.g. the results of a simulation with 500 iterations has a larger probable error than a simulation using 1000 iterations. The error is a function of both the variation in the calculated results and the number of iterations.

During simulations the results can be monitored and the simulations are stopped when a convergence criterion is satisfied. An alternative approach would be to specify the number of iterations and measure the error in the results. This measurement of the error can be done using a technique called "bootstrapping" [AA].

For example, consider the results in table 2, repeated below:

Table 2 The measured and calculated current 90th percentiles for Claremont and Tambo

Number of consumers	Monte Carlo		Calculated		Difference	
	Average [A]	Standard deviation [A]	Average [A]	Standard deviation [A]	Average [A]	Standard deviation [A]
Claremont						
1	16.5	6.1	17.4	6.3	- 0.8	- 0.2
10	119.1	15.8	120.5	15.0	- 1.4	0.7
15	174.7	20.3	176.4	18.1	- 1.7	2.2
Tambo						
1	0.6	0.9	0.9	1.3	- 0.3	- 0.5
10	7.4	2.8	7.6	2.6	- 0.2	0.2
15	11.1	3.2	10.6	2.9	0.4	0.2

The last row of the table compares calculated and simulated results for a section of feeder with 15 consumers connected to one phase. Using boot-strapping, a distribution of the probable range of errors is obtained and shown in the following figure:

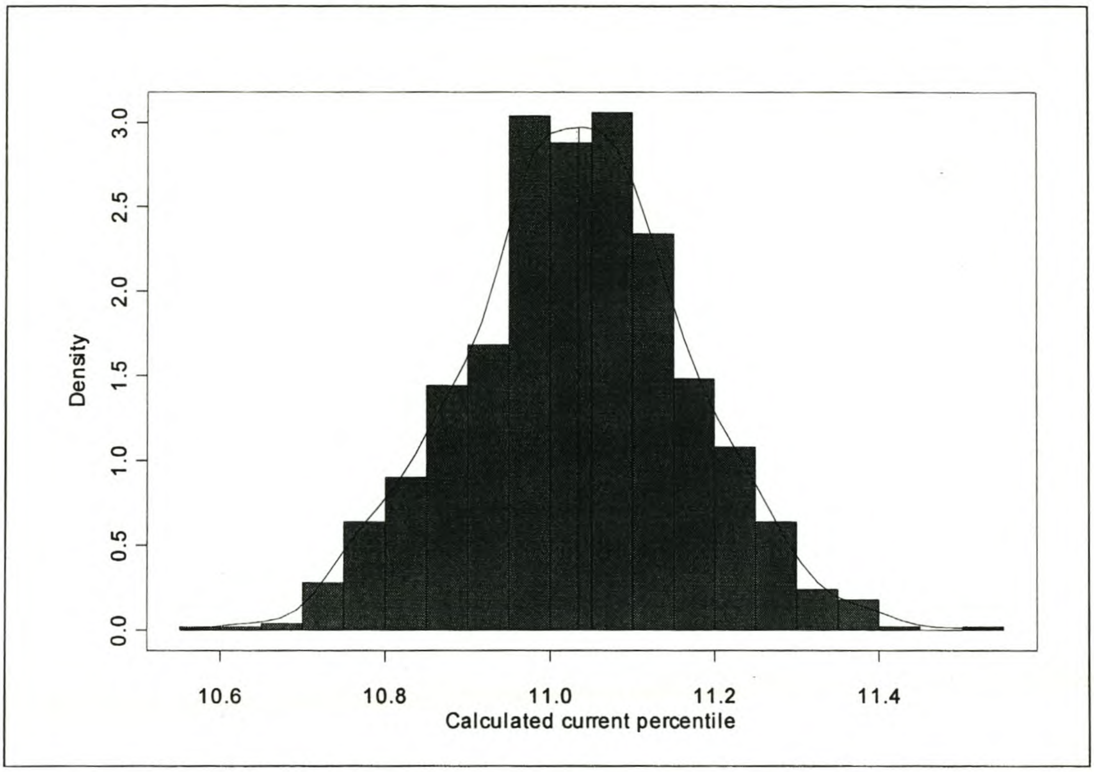

Figure 19 Result of a boot-strapping of the Monte Carlo results obtained for one of the cases reported in table 2.

The calculated value is 10.6 A and this value lies on the edge of the above. It is unlikely that the difference between the calculated and simulated results is due to the Monte Carlo uncertainty only.

Similar results can readily be obtained for any of the Monte Carlo simulations if it is assumed that the error is normally distributed. The error at various levels of confidence can then be calculated using the following equation [43]:

$$Error = \frac{z_\alpha \sigma}{\sqrt{N}}$$

where

Z_α	is the ordinate of the standard normal distribution at some level of confidence
σ	is the simulated standard deviation
N	is the number of iterations

3.10.2 Choice of number of iterations for testing

The choice of number of iterations for the Monte Carlo results based on two considerations:

- A reasonable level of accuracy
- A reasonable time to perform the simulations

The accuracy of the results increases proportionally to the square-root of time to completion. The time per iteration is considerable since each comprises a set of calculations proportional to the period under consideration. For instance, if results are calculated for six month's load data, then a voltage performance calculation is performed for every 5 minute load reading in the period, this amounts to roughly 50 000 voltage drop calculations per iteration. The number of consumers considered in the simulation further increases the complexity and time per voltage drop calculation. The same is true for the resistive loss calculations.

To enable the practical completion of the simulations, the tests were performed with either 500 or 1000 iterations depending on the level of accuracy required. The achieved errors at 95% confidence are stated with the test results.

3.10.3 Other sources of error

The approach followed in all the testing was to estimate the accuracy of the proposed procedures compared with actual measured data. By comparing the results obtained using measured data to the results of the proposed calculation methods, three sources of error may be present:

- The measured load data are from a small numbers of consumers (60-70) in each community, while the proposed load models assumes a continuous distribution of consumers
- Differences between the shape of the theoretical and actual load distributions.
- Errors due to assumptions made to simplify the calculation procedures

Despite these errors, the calculated and simulated results compare well and the calculation procedures are considered accurate enough for practical design purposes.

4. VOLTAGE REGULATION CALCULATION

4.1 INTRODUCTION

A voltage regulation calculation method for residential loads is introduced in this chapter. The method uses the probabilistic load model and the method for calculating the maximum of a linear combination of load currents described in chapter 4 to estimate the voltage regulation on a feeder.

The voltage regulation on a feeder is subject to constraints and these constraints may vary among countries. In South Africa the NRS 048 Quality of supply specification [52] describes a method which is similar to the European specification [22], for assessing the quality of the supply voltage to each consumer. A probabilistic method for designing a feeder for compliance to the NRS 048 criterion is also proposed in this chapter.

A probabilistic method based on the load currents at the moment of system maximum demand is currently the recommended design low voltage feeder design method in South Africa. Some results using the proposed method are compared with results from the recommended method.

4.2 MAXIMUM VOLTAGE DROP

In the previous section it was shown how the mean and variance of the percentiles of a linear combination of current signals could be obtained. The consumer voltage at some point on an LV feeder can be calculated using the following equation:

$$\text{Vcon} = \sqrt{(\text{Vs} - \text{Vre})^2 + \text{Vim}^2} \tag{32}$$

where

V_{re} is the real part of the consumer voltage drop and Vim the imaginary part.

Using a Taylor expansion

$$\text{Vcon} \approx \text{Vs} - \text{Vre} + \frac{\text{Vim}^2}{2\text{Vs}} + \frac{\text{Vre}^2}{2\text{Vs}} \tag{33}$$

If Vim^2 and Vre^2 are small compared to Vs then

$$Vcon \approx Vs - Vre \qquad (34)$$

This assumption will lead to inaccurate calculation of the voltage regulation on the phases not heaviest loaded, in extremely unbalanced cases. In these cases, the assumption leads to an over estimate of the voltage regulation. Since the phase of interest on a low voltage feeder, is the phase with maximum voltage regulation, the over estimate is not considered to be relevant. The methods could however be extended to include additional Taylor expansion terms, but this would increase the complexity and number of additional load parameters.

For a three phase four wire feeder:

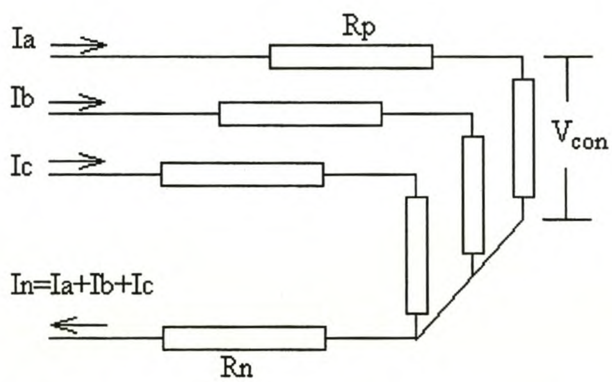

Figure 20 One section of a three phase four wire feeder

Vre is given by

$$Vre = (R_n + R_p)I_a - 0.5R_n(I_b + I_c) \qquad (35)$$

The consumer voltage (Vcon) can then be written as a linear combination of the phase currents, using superposition [30].

Now, for load current trace i, connected to the a-phase on a three phase LV feeder, a_i is calculated as follows:

$$a_i = R_{pi} + R_{ni} \qquad (36)$$

else

$$a_i = -0.5R_{ni} \tag{37}$$

where

R_{pi}	is the phase resistance from the source to consumer i
R_{ni}	is the neutral resistance from the source to consumer i
a_i	is a constant associated with load current i and is used to calculate the voltage regulation based on the principle of superposition

After calculating a_i for each consumer i, the distribution of voltage drop at some level of confidence can be calculated as a linear combination of load traces (see chapter 4).

It should be noted that the MV/LV transformer is not included in the calculations. An equivalent model of the transformer includes series inductance that does not form part of the calculation methods described here. This could be a topic of further research.

4.3 SIMPLIFIED VOLTAGE REGULATION CALCULATION METHOD

The method described above can be very slow if a large number of nodes are present on the low voltage feeder. The method can be significantly simplified if instead of the beta probability distribution, a normal probability distribution is used to estimate the percentile values. This assumption does not imply that the entire load current distribution is normal but rather that the normal distribution can be used to estimate the tail-end distribution of load current. It also does not imply that the limiting distribution of the load current is normal distributed. The assumption is simply that the tail-end of the normal distribution approximates the tail-end of the actual load current distribution with sufficient accuracy for practical design purposes.

In the simplified case the percentile value can be calculated as:

$$V_{\%-tile} = \sum_{i=1}^{N} a_i \mu_i + z_\alpha \sqrt{\sum_{j=1}^{N} a_j^2 \sigma_j^2 + \sum_{k=1}^{N} \sum_{\substack{m=1 \\ k \neq m}}^{N} a_k a_m \sigma_k \sigma_m \rho_{km}} \tag{38}$$

where Z_α is the percentile of the standardized normal distribution at some specified level of confidence.

The expected value of $V_{\%\text{-tile}}$ can be found by replacing σ^2 with the following linear regression on δ:

$$\sigma^2{}_k = G_2 * \delta_k + C_{o2} + e_{k2} \tag{39}$$

where

G_2	is the slope of the relationship between δ_k and σ_k^2
C_{o2}	is the intercept of the relationship between δ_k and σ_k^2
δ_k	is the difference between the current trace k's mean and the mean of the average current trace of all consumers in the group
σ_k^2	is the variance of current trace k
e_{k2}	is the error between the regression line and the actual variance, which is also a distribution with mean = 0 and standard deviation = standard error of the regression

and replacing σ with the following linear regression on δ:

$$\sigma_k = G * \delta_k + C_o + e_k \tag{40}$$

where

G	is the slope of the relationship between δ_k and σ_k
C_o	is the intercept of the relationship between δ_k and σ_k
δ_k	is the difference between the current trace k's mean and the mean of the average current trace of all consumers in the group
σ_k	is the standard deviation of current trace k
e_k	is the error between the regression line and the actual variance, which is also a distribution with mean = 0 and standard deviation = standard error of the regression

Substituting (39) and (40) in (38),

$$V_{\%-tile} = \sum_{i=1}^{N} a_i \mu_i + z_\alpha \sqrt{\begin{array}{l} \sum_{j=1}^{N} a_j^2 (G_2 \delta_j + C_{o2} + e_j) \\ + \sum_{k=1}^{N} \sum_{\substack{m=1 \\ k \neq m}}^{N} a_k a_m (G\delta_k + C_o + e_k)(G\delta_m + C_o + e_m)\rho_{km} \end{array}}$$

If N is large (>10), the effects due to e_j, e_k and e_m are quite small. Typical values can be found in section 3.8. If the standard error of the regression (39) is small, the effects are further reduced. When N is small and the standard error is large, the method will over-estimate the variance of the percentile value.

If C_o and C_{o2} are much greater than the variance of δ then $V_{\%tile}$ can be simplified using a Taylor expansion of the square root:

$$\sqrt{1+x} \approx 1 + \frac{x}{2} \qquad \text{where } |x| << 1 \tag{41}$$

$V_{\%tile}$ reduces to (42)

$$V_{\%-tile} = \sum_{i=1}^{N} a_i \mu_i + z_\alpha \left[\begin{array}{l} \sqrt{K_0} + \frac{1}{2} \sum_{j=1}^{N} \frac{a_j^2 G_2 \delta_j}{\sqrt{K_0}} + \\ \frac{1}{2} \sum_{k=1}^{N} \sum_{\substack{m=1 \\ k \neq m}}^{N} \frac{a_k a_m \rho_{km} (GC_o \delta_k + G.C_o \delta_m + G^2 \delta_k \delta_m)}{\sqrt{K_0}} \end{array} \right]$$

where

$$K_0 = \left[C_{o2} \sum_{j=1}^{N} a_j^2 + C_o^2 \sum_{k=1}^{N} \sum_{\substack{m=1 \\ m \neq k}}^{N} a_k a_m \rho_{km} \right] \tag{43}$$

The expected value of $V_{\%tile}$ (42) and the variance of $V_{\%tile}$ reduce to

$$E[V_{\%tile}] = \mu\sum_{i=1}^{N} a_i + z_\alpha\sqrt{K_1} \tag{44}$$

and

$$V(V_{\%tile}) = V\delta\sum_{k=1}^{N} a_k^2 \cdot \left[1 + \frac{a_k z_\alpha G_2}{2\sqrt{K_1}} + \frac{a_k z_\alpha G C_o}{\sqrt{K_1}} E(\rho)\sum_{\substack{j=1 \\ j\neq k}}^{N} a_j\right]^2 \tag{45}$$

where

G is the slope of the relationship between δ_k and σ_k

μ is the average of the individual load current means

$V\delta$ is the variance of the individual load current means

a_k is calculated for consumer k as

$a_k=R_{nk}+R_{pk}$ if consumer k is connected to the a phase, or

$a_k=-0.5R_{nk}$ if the consumer is connected to either the b or c phase

$R_{nk};R_{pk}$ is the total network resistance between the source and the point of supply of consumer k, for the neutral and phase cables respectively.

Z_α is the percentile value of the standard normal distribution at some level of confidence, α

$E(\rho)$ is the average value of the correlation between individual consumer load currents

K_1 is

$$K_1 = C_{o2}\sum_{j=1}^{N} a_j^2 + C_o^2 E(\rho)\sum_{k=1}^{N}\sum_{\substack{m=1 \\ m\neq k}}^{N} a_k a_m \tag{46}$$

with

C_o is the intercept of the relationship between δ_k and σ_k

This method is simple enough to implement in a spreadsheet, since it does not require the evaluation of any partial derivatives. The method is analytical and very fast. The accuracy in some cases is reduced. These cases specifically are:

- When the number of consumer on the feeder is small (<10) – the simplified algorithm under-estimate the voltage.
- When the variation δ_k is large compared to the intercept of the regression between δ_k and σ_k^2- the simplified algorithm may under or over estimate the voltage.
- When the standard error of the regression between δ_k and σ_k^2 is large – algorithm under-estimates the variance.

Section 4.9 shows a sample implementation of the spreadsheet, with typical values for a poor community in South Africa (Tambo).

4.4 OBTAINING A DESIGN VALUE FROM THE DISTRIBUTION OF PERCENTILE VALUES

The method proposed in section 4.3 estimates the parameters of a distribution of percentile values. The parameters of the distribution are the mean and standard deviation of the voltage performance. A design value can be obtained at a specified level of confidence using the beta distribution, which was accepted at α=0.01 using the Kolmogorov-Smirnov test for goodness of fit [38]. The following steps shows how to calculate the design value at some level of confidence given the mean, $\mu_{\%tile}$ and standard deviation, $\sigma_{\%tile}$ of the voltage performance.

i. Calculate the minimum and maximum values of the linear combination:

$$\text{Maximum} = V_s - \sum_{a_i<0} a_i C_k \tag{47}$$

$$\text{Minimum} = V_s - \sum_{a_i>0} a_i C_k \tag{48}$$

where

C_k	is the circuit breaker size or maximum value of the load current for consumer k
a_i	is defined in section 4.2.
V_s	is the supply voltage (230 V in South Africa)

ii. Scale the average and standard deviation of the voltage performance

$$\mu_{ns} = \frac{\mu_{\%tile} - \text{Minimum}}{\text{Maximum} - \text{Minimum}} \tag{49}$$

$$\sigma_{ns} = \frac{\sigma_{\%tile}}{\text{Maximum} - \text{mininimum}} \tag{50}$$

Note that Minimum = 0 and Maximum = $N.C_k$ with C_k the circuit breaker size

iii. Calculate the α, β using μ_{ns} and σ_{ns}

$$\alpha = \frac{\mu_{ns}{}^2 - \mu_{ns}(\sigma_{ns}^2 + \mu_{ns}^2)}{\sigma_{ns}^2} \tag{51}$$

$$\beta = \frac{\alpha(1 - \mu_{ns})}{\mu_{ns}} \tag{52}$$

The design value can be obtained using the inverse beta function.

$$\hat{V} = \text{qbeta}(\text{cl}, \alpha, \beta)(\text{Maximum} - \text{Minimum}) + \text{Minimum} \tag{53}$$

where

qbeta is the inverse beta distribution function

cl is the confidence level

$\hat{V}$ is the design percentile at some level of confidence

Section 4.11 shows an implementation in Mathcad 2000 of the above procedure.

4.5 SPECIFICATION OF VOLTAGE COMPATIBILITY LEVELS

The minimum voltage performance of a low voltage feeder is usually specified with a maximum and minimum value for the voltage. The IEC specification on standard voltages [35] specifies that the voltage at the point of supply will be within ±10% of the nominal value

(230V). A similar document is produced by ANSI and it specifies two ranges of voltage limits, as follows:

Table 3 Voltage ranges according to ANSI C84.1

Range	Voltage range	Restriction
A	±5%	Voltage should stay within this range most of the time
B	- 8.3% / +5.83%	Corrective action should be taken if voltage exceeds these limits

Neither of these two documents specifies how the quality of the voltage will be assessed. A Euro Norm document, however, exists which specifies the assessment method [22]. This document is in the process of being accepted in many European countries as national standards. The voltage performance method will be referred to in this text as the EN50160 voltage performance criterion or EN50160 voltage performance. The specification reads as follows:

"2.3 Supply voltage variations

Under normal operating conditions, excluding situations arising from faults or voltage interruptions,

- *During each period of one week 955 of the 10 min mean rms values of the supply voltage shall be winthin the range U_n ±10%*
- *All 10 minute mean rms values of the supply voltage shall be within the range U_n +10% / -15%"*

In South Africa the voltage regulation assessment method for low voltage feeder based on measurement is specified by [52] and will in this text be referred to as the NRS 048 voltage performance criterion or the voltage performance of a feeder. The specification is as follows:

"4.6.2 Assessment method

The assessment period is a minimum of 7 continuous days.
On each phase of the supply voltage, for each 24h day (00:00 to 24:00), the highest 10 min root-mean-square values of the supply voltage within which the voltage remains for 95% of the time are recorded for each phase and the highest of these is retained as a daily value.
Similarly, the lowest 10 min root-mean-square values of the supply voltage within which the voltage remains for 95% of the time are recorded for each phase and the lowest of these is retained as a daily value.
NB - Measurements shall be taken at the extremities (near and far ends) of feeders
All incidents where two or more consecutive 10 min values are outside the compatibility levels shall be recorded."

The voltage performance in NRS 048 and EN 50160 voltage performance can be approximated as a percentile of the voltage drop distribution of a group of consumers as calculated using the methods described in section 4.2. The calibration of the percentile value for the first constraint in EN50160 voltage performance criterion has not been done and could be the topic of further research. The second constraint is simply the maximum value over a period and can be calculated as:

$$\mathrm{cl} = \frac{\mathrm{Period} - 1}{\mathrm{Period}} \tag{25}$$

where

cl is the confidence level for which the percentile distribution should be calculated

Period is the period of evaluation expressed as number of 10 minutes

For the NRS 048 voltage performance, the confidence level for the percentiles can be estimated using the following equation:

$$\mathrm{cl} = \frac{\mathrm{Period} - \mathrm{days} * 7}{\mathrm{Period}} \tag{25}$$

where

cl	is the confidence level for which the percentile distribution should be calculated
Period	is the period of evaluation expressed as number of 10 minutes
"Days"	is the number of days on which violations is permitted to occur (according to NRS 048 this is 1)

The calculated voltage performance using the confidence level as calculated above is not the same as the NRS 048 voltage performance. The difference is that

- the calculation procedure obtains a description of the load current distribution for a period of time and then estimates a percentile value for the entire period.
- the NRS 048 voltage performance is evaluated every day and the 95th percentile of the load currents for each day is recorded. The worst value of these recorded voltage values is then taken to be the voltage performance of the feeder.

Equation 25 is based on the assumption that the seven highest ten minute average load currents occur on the day with the worst voltage performance and the seven second highest ten minute average load currents occur on the day with the second worst voltage performance etc. This assumption may cause the voltage performance calculation procedure to under or overestimate the voltage performance depending on the actual distribution of the ten-minute average load currents on the different days with high load.

Tests performed using the load data from Tambo and Claremont indicate that this assumption gives results that is suitable for practical design purposes. The refinement of methods for the selection of the confidence level for different types of communities could be the topic of future research.

The number of days with voltage violations is given as a variable in equation 25 since it is assumed that not all violations will be either detected or considered when the voltage performance of a feeder is assessed. The specific implementation in South Africa and how it will be enforced was not specified when this document was written.

The NRS 048 and EN 50160 voltage regulation criteria further calls for the load readings to be RMS values sampled at 3s intervals and averaged for 10 minute periods. The load readings used to test the methods described here were 5 minute averaged readings. The relationship between these RMS readings and the load research data used to compile these methods could not be evaluated. This could be a topic of future research.

4.6 TESTING

The methods described in this section were tested with data from a low income community using Monte Carlo simulations. The simulations took more than 144 hours to complete on three PC's with Pentium or better and 64 MB or more memory. Each process reported results individually and the reports were combined after the simulations were completed.

The Monte Carlo is summarized as follows:

1. Given: a network with N consumers
2. Select at random *with replacement* N consumers from a community in the NRS load research database
3. For the specified period calculate for each 5-minute interval the voltage drop, loss, energy consumption and maximum current. Loads are modelled as constant current to simplify the calculations
4. Determine the voltage regulation performance using the technique as outlined in NRS 048
5. Repeat steps 2-4 a thousand times and store the result after each iteration

The voltage performance evaluation method was tested on a low income community (Tambo) for five months (April – August, inclusive).

4.6.1 Balanced feeder

The NRS 048 voltage performance calculation method was tested for a low income community (Tambo) using the following network:

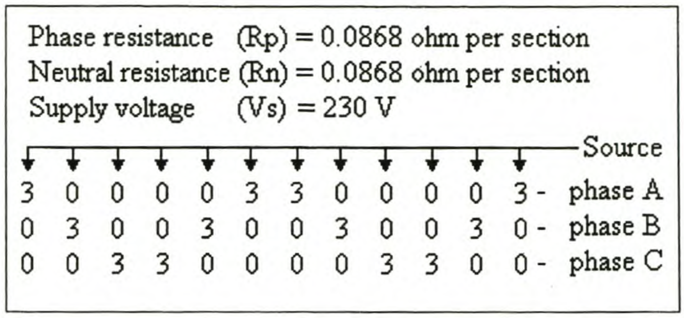

Figure 21 Balanced three phase feeder used to test the NRS 048 voltage performance calculation procedure

The network is summarized in the following table:

Table 4 A balanced network used to test the NRS 048 voltage performance calculation method

	Consumer connections			Conductor resistance (Ω)	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1	3			0.0868	0.0868
2		3		0.0868	0.0868
3			3	0.0868	0.0868
4			3	0.0868	0.0868
5		3		0.0868	0.0868
6	3			0.0868	0.0868
7	3			0.0868	0.0868
8		3		0.0868	0.0868
9			3	0.0868	0.0868
10			3	0.0868	0.0868
11		3		0.0868	0.0868
12	3			0.0868	0.0868

Column ma is the phase under being consideration, column mb and mc is the other two phases.

The method was tested using various numbers of days with violations and the results are shown in the following table:

Table 5 A comparison between the simulated and predicted NRS 048 voltage performance of the balanced three phase feeder

Days	Simulated			Calculated		Difference	
	Average [V]	Standard deviation [V]	Error at 90% confidence	Average [V]	Standard deviation [V]	Average [V]	Standard deviation [V]
1.0	210.2	7.6	0.4	212.5	6.7	- 2.3	0.8
2.0	212.2	6.7	0.3	213.3	6.5	- 1.1	0.2
3.0	213.1	6.4	0.3	213.9	6.3	- 0.7	0.1
4.0	213.7	6.3	0.3	214.2	6.2	- 0.5	0.1
5.0	214.2	6.2	0.3	214.5	6.1	- 0.4	0.1
6.0	214.5	6.1	0.3	214.8	6.0	- 0.2	0.1
7.0	214.8	6.1	0.3	215.0	5.9	- 0.1	0.1
8.0	215.2	6.0	0.3	215.2	5.9	- 0.0	0.1
9.0	215.4	5.9	0.3	215.3	5.8	0.1	0.1
10.0	215.7	5.9	0.3	215.5	5.8	0.2	0.1
20.0	217.2	5.5	0.3	216.5	5.4	0.7	0.1

The difference between the measured and simulated results has a standard error of 0.4 V and the maximum error associated with random selection (Monte Carlo) at 90% confidence was also 0.4 V. When the voltage performance is calculated with "number of days with violations" equal to 1, the difference between the calculated and simulated voltage performance is large. It is clear that a strong relationship exists between the error and the "number of days with violations". This relationship could be due to :

- The simulation method – the voltage performance for all the "number of days with violations" were calculated using the same iteration of the Monte Carlo simulation. A systematic error is therefore introduced in the results due to the simulation method.
- The difference between the actual distribution of periods with extremely high load and the theoretical distribution of load current (see section 4.5)

Section 4.9 shows the parameters and calculations for this feeder.

4.6.2 Single phase feeder

The NRS 048 voltage performance calculation method was also tested for a poor community (Tambo) using the following feeder:

Phase resistance (Rp) = 0.0434 ohm per section
Neutral resistance (Rn) = 0.0434 ohm per section
Supply voltage (Vs) = 230 V

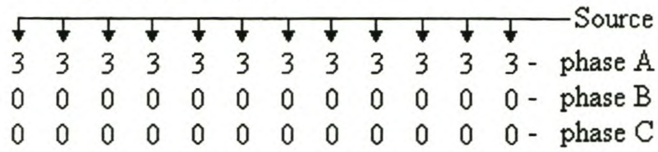

Figure 22 Single phase feeder used to test the NRS 048 voltage performance calculation method

Table 6 A summary of the single phase feeder used to test the NRS 048 voltage performance calculation method

	Consumer connections			**Conductor resistance per section (Ω)**	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1	3			0.0434	0.0434
2	3			0.0434	0.0434
3	3			0.0434	0.0434
4	3			0.0434	0.0434
5	3			0.0434	0.0434
6	3			0.0434	0.0434
7	3			0.0434	0.0434
8	3			0.0434	0.0434
9	3			0.0434	0.0434
10	3			0.0434	0.0434
11	3			0.0434	0.0434
12	3			0.0434	0.0434

The following table compares simulated and calculated results for the network specified in the table above.

Table 7 A comparison between the simulated and predicted NRS 048 voltage performance of a single phase feeder

Days	Simulated			Calculated		Difference	
	Average [V]	Standard deviation [V]	Error at 90% confidence	Average [V]	Standard deviation [V]	Average [V]	Standard deviation [V]
1.0	211.4	4.5	0.2	211.9	4.6	0.5	0.1
5.0	214.2	3.8	0.2	214.1	4.0	- 0.1	0.2
10.0	215.9	3.4	0.2	215.4	3.7	- 0.5	0.3
15.0	217.0	3.3	0.2	216.1	3.5	- 0.9	0.2

The maximum error between the simulated and calculated values is 0.9 V. The maximum error associated with random sampling (Monte Carlo) at 90% confidence is 0.2 V. The remainder of the error could be due to the differences between the assumed load distributions and the actual distributions. A further source of error is the small number of consumers sampled from the community resulting in a discreet distribution and not a continuous distribution as assumed (see section 3.10 for general discussion).

Section 4.10 shows the parameters and calculations for this feeder.

4.6.3 Short three phase feeder

The calculation method was further tested for a balanced three phase feeder with load data from a medium income community (Claremont). The following table shows a specification of the test feeder:

Table 8 A short three phase feeder used to test the NRS 048 voltage performance calculation method

	Consumer connections			Conductor resistance per section (Ω)	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1	1	1	1	0.0868	0.0868
2	1	1	1	0.0868	0.0868
3	1	1	1	0.0868	0.0868

The following table compares the simulated and calculated voltage performance for the feeder:

Table 9 A comparison between the simulated and predicted NRS 048 voltage performance of a balanced three-phase feeder

Days	Simulated			Calculated		Difference	
	Average [V]	Standard deviation [V]	Error at 90% confidence	Average [V]	Standard deviation [V]	Average [V]	Standard deviation [V]
1.0	213.5	4.6	0.2	212.3	4.5	1.3	0.1
5.0	216.4	4.1	0.2	214.8	4.1	1.6	- 0.1
10.0	217.7	3.9	0.2	216.1	3.9	1.6	- 0.1
15.0	218.6	3.8	0.2	217.4	3.8	1.2	0.0

The simulated and calculated results show some bigger differences. This could be due to the small number of consumers on the feeder or the small number of Monte Carlo simulations. A further source of error could be the difference between the actual distribution of periods with extremely high load and the theoretical distribution of load current (see section 4.5 and section 3.10 for a general discussion)

4.7 COMPARISON OF NRS 048 VOLTAGE PERFORMANCE METHOD AND NRS 034 RECOMMENDED METHOD

The minimum voltage performance of a low voltage residential feeder in South Africa is specified in NRS 048 (see above). NRS 034 contains a recommended voltage calculation method, which is called the Herman Beta method [51]. This method was developed by Herman [29] and later improved on by Herman and Heunis [30].

The Herman Beta method is a probabilistic method and uses the loads at the time of system maximum demand to calculate a probability distribution of the voltages at system maximum demand using a moments. A confidence interval is specified and a percentile value is calculated after a beta probability distribution has been fitted.

Some of the properties and assumptions in the Herman Beta method and proposed calculation method are compared in the following table.

Table 10 A comparison of the Herman Beta method and the proposed calculation method based on NRS048

Properties and assumptions	Herman Beta method	Proposed calculation method (NRS 048)
Load dependence	Independent	Dependent and correlated
Period of consideration	A single of instance of high demand	An entire period of time
Load distribution	Beta probability distribution of load currents at an instant in time	Beta probability distribution for the load current of each distribution

In the following table the calculated voltage performance at a 10% level of confidence using the Herman Beta method and the method proposed in this section based on NRS 048 is compared. All calculations were performed with data from a poor community (Tambo).

Calculated parameter	Value
Feeder in section 4.6.1	
$V_{\%tile}$ at 90% confidence	203.83
Herman beta design value at 90% confidence	202.32
Feeder in section 4.6.2	
$V_{\%tile}$ at 90% confidence	205.8
Herman beta design value at 90% confidence	185.93

The Herman Beta parameters used for the poor community (Tambo) was α=0.16, β=2.06 and Circuit breaker = 20 A.

A comparison of the two method with data from a medium income community (Claremont) is shown in the following table:

Calculated parameter	**Value**
Feeder in section 4.6.1	
$V_{\%tile}$ at 90% confidence	206.50
Herman beta design value at 90% confidence	212.51

The Herman Beta parameters used for the medium income community (Claremont) was α=1.91, β=4.92 and circuit breaker = 60A.

It seems that the Herman Beta method could:

- Over-estimate the voltage drop
- accurately estimate the voltage drop
- under-estimate the voltage drop

The confidence level of the Herman Beta design method can be adjusted to always accurately estimate the voltage drop. The calibration of the confidence level is a topic for further research.

4.8 CONCLUSION AND SUMMARY

A probabilistic voltage regulation calculation method is described in this chapter. The method can be used to estimate the NRS 048 voltage performance of a low voltage residential feeder. The method was tested using a Monte Carlo type simulation with actual load data from a poor and medium income community. The calculated and simulated results compare well.

The method is analytical, fast and simple enough to implement in a spreadsheet.

4.9 APPENDIX: CALCULATIONS FOR BALANCED THREE PHASE NETWORK

This section contains sample input parameters, calculations and results for the three phase network used to test the voltage calculation method. Although NRS 048 specifies that the number of days with violations should be 1, the number of days with violations is taken to be 10 for this sample calculation.

4.9.1 Input parameters:

$N = 36$ $c_k = 20$ $e_o = 0$

$\mu = 0.322$

$V\delta = 0.15$ $Ve_o = 0.366^2$ $days = 10$

$G = 1.41$ $C_o = 0.63$ $\rho = 0.07$

$G2 = 3.89$ $C_{o2} = 0.82$ $V_s = 230$

$\delta = 0$ $e_o = 0$

$z_\alpha = \mathrm{qnorm}(cl, 0, 1)$ $cl = \dfrac{(180 \cdot 24 \cdot 6 - days \cdot 7)}{180 \cdot 24 \cdot 6}$

4.9.2 Calculations and results:

- $$K_1 = C_{o2} \cdot \sum_{j=0}^{N-1} \left(a_j\right)^2 + C_o{}^2 \cdot \rho \cdot \left(\sum_{k=0}^{N-1} a_k\right)^2 - C_o{}^2 \cdot \rho \cdot \left[\sum_{m=0}^{N-1} \left(a_m\right)^2\right]$$

$K_1 = 19.62664$

$$\text{E(V\%tile)} = V_s - \left(\mu \cdot \sum_{i=0}^{N-1} a_i + z_\alpha \cdot \sqrt{K_1}\right)$$

E(V%tile)= 215.495

$$\text{V(V\%tile)} = V\delta \cdot \sum_{k=0}^{N-1} \left(a_k\right)^2 \cdot \left(1 + \frac{a_k \cdot z_\alpha \cdot G2}{2 \cdot \sqrt{K_1}} + \frac{z_\alpha \cdot G C_o \cdot \rho}{2 \cdot \sqrt{K_1}} \cdot \sum_{m=0}^{N-1} a_m - \frac{z_\alpha \cdot G C_o \cdot \rho}{2 \cdot \sqrt{K_1}} \cdot a_k\right)^2$$

$\sqrt{\text{V(V\%tile)}} = 5.76516$

4.9.3 Values for a_i

i	a_i	i	a_i	i	a_i
1	0.174	2	0.174	3	0.174
4	-0.087	5	-0.087	6	-0.087
7	-0.13	8	-0.13	9	-0.13
10	-0.174	11	-0.174	12	-0.174
13	-0.217	14	-0.217	15	-0.217
16	1.042	17	1.042	18	1.042
19	1.215	20	1.215	21	1.215
22	-0.347	23	-0.347	24	-0.347
25	-0.391	26	-0.391	27	-0.391
28	-0.434	29	-0.434	30	-0.434
31	-0.477	32	-0.477	33	-0.477
34	2.083	35	2.083	36	2.083

4.10 APPENDIX: CALCULATIONS FOR SINGLE PHASE NETWORK

This section shows the input parameters and calculations for the single-phase network used to test the voltage performance method. Although NRS 048 specifies that the number of days with violations should be 1, the number of days with violations is taken to be 10 for this sample calculation.

4.10.1 Input parameters

$N = 36$ $c_k = 20$ $e_o = 0$

$\mu = 0.28$

$V\delta = 0.15$ $Ve_o = 0.366^2$ $days = 15$

$G = 1.54$ $C_o = 0.5$ $\rho = 0.08$ $V_s = 230$

$G2 = 4.55$ $C_{o2} = 0.7$

$\delta = 0$ $e_o = 0$ $cl = \frac{(27 \cdot 24 \cdot 6 - 7 \cdot days)}{27 \cdot 24 \cdot 6}$ $z_\alpha = \text{qnorm}(cl, 0, 1)$

4.10.2 Calculations

$$K_1 = C_{o2}\cdot\sum_{j=0}^{N-1}\left(a_j\right)^2 + C_o{}^2\cdot\rho\cdot\left(\sum_{k=0}^{N-1} a_k\right)^2 - C_o{}^2\cdot\rho\cdot\left[\sum_{m=0}^{N-1}\left(a_m\right)^2\right]$$

$$K_1 = 18.24453$$

$$E(V\%tile) = V_s - \left(\mu\cdot\sum_{i=0}^{N-1} a_i + z_\alpha\cdot\sqrt{K_1}\right)$$

$$E(V\%tile) = 216.083$$

$$V(V\%tile) = V\delta\cdot\sum_{k=0}^{N-1}\left(a_k\right)^2\cdot\left(1 + \frac{a_k\cdot z_\alpha\cdot G2}{2\cdot\sqrt{K_1}} + \frac{z_\alpha\cdot G\cdot C_o\cdot\rho}{\sqrt{K_1}}\cdot\sum_{m=0}^{N-1} a_m - \frac{z_\alpha\cdot G\cdot C_o\cdot\rho}{\sqrt{K_1}}\cdot a_k\right)^2$$

$$\sqrt{V(V\%tile)} = 3.5407$$

4.10.3 Values for a_i

i	a_i	i	a_i	i	a_i
1	0.087	2	0.087	3	0.087
4	0.174	5	0.174	6	0.174
7	0.26	8	0.26	9	0.26
10	0.347	11	0.347	12	0.347
13	0.434	14	0.434	15	0.434
16	0.521	17	0.521	18	0.521
19	0.608	20	0.608	21	0.608
22	0.694	23	0.694	24	0.694
25	0.781	26	0.781	27	0.781
28	0.868	29	0.868	30	0.868
31	0.955	32	0.955	33	0.955
34	1.042	35	1.042	36	1.042

4.11 APPENDIX: CALCULATIONS FOR SHORT THREE PHASE NETWORK AND PERCENTILE VALUE OF VOLTAGE PERFORMANCE DISTRBUTION

This section shows the calculation for a short three phase network (section 4.6.3). The method for calculating the percentile value of the voltage performance distribution is also shown.

$N = 9$ $c_k = 60$ $e_o = 0$

$\mu = 7.28$

$V\delta = 13.77$ $Ve_o = 0.366^2$ $days = 1$

$G2 = 5.81$ $C_{o2} = 50.56$ $\rho = 0.18$

$C_o = 6.887$

$G = 0.39$

$\delta = 0$ $e_o = 0$ $cl = \dfrac{(31 \cdot 24 \cdot 6 - days \cdot 7)}{31 \cdot 24 \cdot 6}$

$V_s = 230$

$z_\alpha = \text{qnorm}(cl, 0, 1)$

4.11.1 Values for a_i

i	a_i	i	a_i	i	a_i
1	0.174	2	-0.043	3	-0.043
4	0.347	5	-0.087	6	-0.087
7	0.521	8	-0.13	9	-0.13

$$K_1 = C_{o2}\cdot\sum_{j=0}^{N-1}\left(a_j\right)^2 + C_o^{\,2}\cdot\rho\cdot\left(\sum_{k=0}^{N-1} a_k\right)^2 - C_o^{\,2}\cdot\rho\cdot\left[\sum_{m=0}^{N-1}\left(a_m\right)^2\right]$$

$K_1 = 22.27701$

$$E(V\%tile) = V_s - \left(\mu\cdot\sum_{i=0}^{N-1} a_i + z_\alpha\cdot\sqrt{K_1}\right)$$

$E(V\%tile) = 212.257$

$$V(V\%tile) = V\delta\cdot\sum_{k=0}^{N-1}\left(a_k\right)^2\cdot\left(1 + \frac{a_k\cdot z_\alpha\cdot G2}{2\cdot\sqrt{K_1}} + \frac{z_\alpha\cdot G\cdot C_o\cdot\rho}{2\cdot\sqrt{K_1}}\cdot\sum_{m=0}^{N-1} a_m - \frac{z_\alpha\cdot G\cdot C_o\cdot\rho}{2\cdot\sqrt{K_1}}\cdot a_k\right)^2$$

$\sqrt{V(V\%tile)} = 4.48226$

$Minimum = V_s - \sum matrix(N, 1, MaximumCalc)\cdot c_k$ $\qquad$ $Minimum = 167.48$

$Maximum = V_s - \sum matrix(N, 1, MinimumCalc)\cdot c_k$ $\qquad$ $Maximum = 261.2$

$$\mu_{ns} = \frac{(E(V\%tile) - Minimum)}{Maximum - Minimum}$$

$\mu_{ns} = 0.478$

$$\sigma_{ns} = \left[\frac{\sqrt{V(V\%tile)}}{(Maximum - Minimum)}\right]$$

$\sigma_{ns}^{\,2} = 2.287\times 10^{-3}$

$$\alpha = \frac{\left[\mu_{ns}^{\,2} - \left[\left(\sigma_{ns}^{\,2} + \mu_{ns}^{\,2}\right)\cdot\mu_{ns}\right.\right.}{\sigma_{ns}^{\,2}}$$

$$\beta = \frac{\alpha\left(1 - \mu_{ns}\right)}{\mu_{ns}}$$

$\alpha = 51.639$

$cl = 0.1$

$\beta = 56.443$

$$V_{design} = qbeta\left(cl, \alpha, \beta\right)\cdot(Maximum - Minimum) + Minimum$$

$V_{design} = 206.503$

5. RESISTIVE LOSS CALCULATION

5.1 INTRODUCTION

The loss on a feeder assume a range of values due to the range of different consumers which might be connected to it. The range can be expressed as a probability distribution if enough information is available about the consumers connected to it.

This section shows how the expected value and the variance of the loss on an low voltage feeder can be calculated. The loss is calculated for some period using the load model described in section 3.2. A percentile value for the loss distribution can be obtained by fitting a beta probability distribution to the minimum, maximum, expected value and variance of the range of loss values.

5.2 RESISTIVE LOSS FOR ONE SECTION

The technical loss on a section of a feeder can be calculated as:

$$\text{Loss} = I_a^2R_p+I_b^2R_p+I_c^2R_p+R_nI_n^2 \tag{54}$$

where

I_a	is the current in the a phase of the section
I_b	is the current in the b phase of the section
I_c	is the current in the c phase of the section
I_n	is the current in the neutral conductor of the section
R_p	is the phase resistance
R_n	is the neutral resistance

The average value of the loss for a period of time is calculated as follows:

$$\mu_{loss} = E[I_a^2]R_p+E[I_b^2]R_p+E[I_c^2]R_p+R_n\, E[I_n^2] \tag{55}$$

The technical loss in a feeder can be calculated over a period of time using the following property of random variables:

$$E[I^2] = E[I]^2 + V[I] \tag{56}$$

where

$E[I^2]$ is the second moment around the origin

$E[I]$ is the average of the random variable

$V[I]$ is the variance of the random variable

The average loss is therefore

$$\mu_{loss} = E[I_a]^2 R_p + E[I_b]^2 R_p + E[I_c]^2 R_p + E[I_n]^2 R_n + V[I_a] R_p + V[I_b] R_p + V[I_c] R_p + V[I_n] R_n \quad (57)$$

I_n can be written as

$$I^2_n = (I_a - 0.5(I_b + I_c))^2 + 0.75(I_b + I_c)^2 \quad (58)$$

The average loss becomes

$$\mu_{loss} = (E[I_a]^2 + E[I_b]^2 + E[I_c]^2 + V[I_a] + V[I_b] + V[I_c])(R_p + R_n) - R_n (E[I_a I_b] + E[I_b I_c] + E[I_b I_c]) \quad (59)$$

The following properties of random variables are useful to simplify the equation

$$E[X.Y] = E[X]E[Y] + \rho_{xy}\sigma_x\sigma_y \quad (60)$$

$$E[X^2] = E[X]^2 + \rho\sigma_x^2 \quad (61)$$

And substituting the following

$$E[I] = \sum_{k=1}^{N} \mu_k \quad (62)$$

$$V[I] = \sum_{k=1}^{N} \sigma_k^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \rho_{ij}\sigma_i\sigma_j \quad (63)$$

yields the following equation for the average loss

$$\mu_{loss} = (E[I_a^2] + E[I_b^2] + E[I_c^2])(R_p + R_n) - (E[I_a I_b] + E[I_b I_c] + E[I_a I_c])R_n \qquad (64)$$

where

$$E[I_a^2] = \sum_{k=1}^{N} (\mu_k^2 + \sigma_k^2) Q_{ak} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j) Q_{ai} Q_{aj} \qquad (65)$$

$$E[I_b^2] = \sum_{k=1}^{N} (\mu_k^2 + \sigma_k^2) Q_{bk} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j) Q_{bi} Q_{bj} \qquad (66)$$

$$E[I_c^2] = \sum_{k=1}^{N} (\mu_k^2 + \sigma_k^2) Q_{ck} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j) Q_{ci} Q_{cj} \qquad (67)$$

$$E[I_a I_b] = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j) Q_{ai} Q_{bj} \qquad (68)$$

$$E[I_a I_c] = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j) Q_{ai} Q_{cj} \qquad (69)$$

$$E[I_b I_c] = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (\mu_i \mu_j + \rho_{ij} \sigma_i \sigma_j) Q_{bi} Q_{cj} \qquad (70)$$

N	is the total number of consumers supplied by the section (N_a+N_b+N_c)
Q_{ak}	is 1 if consumer k is supplied by phase A of the section else 0
Q_{bk}	is 1 if consumer k is supplied by phase B of the section else 0
Q_{ck}	is 1 if consumer k is supplied by phase C of the section else 0
μ_{loss}	is the average loss for a period of time
R_p	is the phase resistance
R_n	is the neutral resistance
μ_k	is the average load current of consumer k

σ_k is the standard deviation of the load current of consumer k

ρ_{ij} is the correlation between load current i and load current j

From the load model in section 3.2:

$$\mu_k = \mu + \delta_k \tag{71}$$

$$\sigma^2{}_k = G_2 * \delta_k + C_{o2} \tag{72}$$

$$\sigma_k = G * \delta_k + C_o \tag{73}$$

Where

μ_k is the average of the current trace for consumer k

μ is the expected value of the average current traces for a group of consumers

G_2 is the slope of the relationship between δ_k and σ_k^2

C_{o2} is the intercept of the relationship between δ_k and σ_k^2

δ_k is the difference between the current trace k's mean and the mean of the average current trace of all consumers in the group

σ_k^2 is the variance of current trace k

G is the slope of the relationship between δ_k and σ_k

Co is the intercept of the relationship between δ_k and σ_k

σ_k is the standard deviation of current trace k

Note that the error terms of above regressions are ignored, since they are considered small.

The final set of equations for the average loss for a specific section is obtained by replacing the load model equations into (64).

Average loss equations:

$$\mu_{loss} = (E[I_a^2] + E[I_b^2] + E[I_c^2])(R_p + R_n) - (E[I_a I_b] + E[I_b I_c] + E[I_a I_c])R_n \tag{74}$$

where

$$E[I_a^2] = \sum_{k=1}^{N} J_k Q_{ak} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} P_{ij} Q_{ai} Q_{aj} \tag{75}$$

$$E[I_b^2] = \sum_{k=1}^{N} J_k Q_{bk} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} P_{ij} Q_{bi} Q_{bj} \tag{76}$$

$$E[I_c^2] = \sum_{k=1}^{N} J_k Q_{ck} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} P_{ij} Q_{ci} Q_{cj} \tag{77}$$

$$E[I_a I_b] = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (P_{ij}) Q_{ai} Q_{bj} \tag{78}$$

$$E[I_a I_c] = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (P_{ij}) Q_{ai} Q_{cj} \tag{79}$$

$$E[I_c I_b] = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} (P_{ij}) Q_{ci} Q_{bj} \tag{80}$$

with

$$K_1 = \mu^2 + C_{o2} \tag{81}$$

$$K_2 = 2\mu + G_2 \tag{82}$$

$$K_{3ij} = \mu^2 + \rho_{ij} C_o^2 \tag{83}$$

$$K_{4ij} = \mu + \rho_{ij} G C_o \tag{84}$$

$$K_{5ij} = 1 + \rho_{ij} G^2 \tag{85}$$

$$J_k = (K_1 + K_2 \delta_k + \delta_k^2) \tag{86}$$

$$P_{ij} = K_{3ij} + K_{4ij}(\delta_i + \delta_j) + K_{5ij} \delta_i \delta_j \tag{87}$$

N is the total number of consumers supplied by the section (N_a+N_b+N_c)

Q_{ak} is 1 if consumer k is supplied by phase A of the section else 0

Q_{bk} is 1 if consumer k is supplied by phase B of the section else 0

Q_{ck} is 1 if consumer k is supplied by phase C of the section else 0

μ_{loss} is the average loss for a period of time

R_p is the phase resistance

R_n is the neutral resistance

5.3 DISTRIBUTION OF TECHNICAL LOSS

A general equation for the average technical loss for N consumers supplied from a section was derived in the previous section. The loss is expressed in terms of the μ_k and σ_k, the average and standard deviation of load current k. Since μ_k and σ_k can assume a range of possible values, the calculated loss can also assume a range of possible values.

In this section it is assumed that enough information is available about the load currents in a community to be able to calculate a probability distribution for the losses on a feeder.

The expected value of the average loss can be calculated using the following steps:

i. Calculate the expected value of the following coefficients:

$$E(J_k) = K_1 + E(\delta^2) \tag{88}$$

$$E(P_{ij}) = E(K_{3ij}) \tag{89}$$

$$K_1 = \mu^2 + C_{o2} \tag{90}$$

$$E(K_{3ij}) = \mu^2 + E(\rho_{ij})C_o^2 \tag{91}$$

where

$E(\rho_{ij})$ is the expected value of the correlation coefficient and is assumed independent of the standard deviations of the load currents of consumers i and j.

$E(\delta^2)$ is the variance of the individual consumer current means and is equal to $V\delta$

Note that some of the coefficients are not shown since they reduce to zero ($E(\delta_k)=0$).

ii. Calculate Q_a,Q_b and Q_c

Q_a, Q_b and Q_c are used to indicate the presence or absence of a consumer load current at a specific node. A convenient way of doing this is to regard Q_a as a connection matrix, where each of the columns is a different node and each of the rows indicate the presence or absence of a specific load current. The presence is indicated as 1 and the absence is indicated as 0.

For example:

A feeder is specified in the following table:

Table 11 A short three phase network used to illustrate how Q_a, Q_b and Q_c is obtained

	Consumer connections			Conductor resistance per section (Ω)	
Node	**Ma**	**Mb**	**Mc**	**Rp**	**Rn**
1	1	1	1	0.0868	0.0868
2	1	1	1	0.0868	0.0868
3	1	1	1	0.0868	0.0868

Q_a, Q_b and Q_c is given below:

Table 12 Qa for the feeder specified above

Consumer	Node 1	Node 2	Node 3
1	1	1	1
2	1	1	0
3	1	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0

Table 13 Qb for the feeder specified above

Consumer	Node 1	Node 2	Node 3
1	0	0	0
2	0	0	0
3	0	0	0
4	1	1	1
5	1	1	0
6	1	0	0
7	0	0	0
8	0	0	0
9	0	0	0

Table 14 Qc for the feeder specified above

Consumer	Node 1	Node 2	Node 3
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	1	1	1
8	1	1	0
9	1	0	0

iii. Calculate the expected value of the average loss

Equation (92):

$$E[\mu_{loss}] = \sum_{s=1}^{Sections}\left[\begin{array}{l} \sum_{k=1}^{N} E[J_k](Qa_{s,k} + Qb_{s,k} + Qc_{s,k})(Rp_s + Rn_s) \\ + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N} E[P_{ij}](Qa_{s,i}Qa_{s,j} + Qb_{s,i}Qb_{s,j} + Qc_{s,i}Qc_{s,j})(Rn_s + Rp_s) \\ - \sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N} E[P_{ij}](Qa_{s,i}Qb_{s,j} + Qb_{s,i}Qc_{s,j} + Qb_{s,i}Qc_{s,j})Rn_s \end{array}\right]$$

where

Rn_s	is the neutral resistance of the cable in section s
Rp_s	is the phase resistance of the cable in section s
$Qa_{s,k}$	is the presence (1) or absence (0) of load current k in section s on phase a
$Qb_{s,k}$	is the presence (1) or absence (0) of load current k in section s on phase b
$Qc_{s,k}$	is the presence (1) or absence (0) of load current k in section s on phase c

The variance of the average loss on a feeder can be calculated using the following steps based on a second order Taylor approximation :

$$V[\mu_{loss}] \approx \sum_{k=1}^{N} T1_k^2 V\delta + \frac{T2_k^2}{4} V\delta 2 + T1_k T2_k E[\delta_k^3] \tag{93}$$

where

$V[\mu_{loss}]$	is the variance of the average loss
$T1_k^2$	is the value of the first derivative of the average loss to δ_k with $\delta_k=0$
$T2_k^2$	is the value of the second derivative of the average loss to δ_k with $\delta_k=0$
$V\delta$	is the variance of the individual load current means
$V\delta 2$	is the variance of the squares of the individual load current means
$E[\delta^3]$	is the third moment of the individual load current means
N	is the total number of consumers

iv. Calculate the first and second order partial derivatives of the coefficients of (64) with δ_k=0

$$\left.\frac{\partial J_k}{\partial \delta_k}\right|_{(\delta_k=0)} = K_2 \tag{94}$$

$$\left.\frac{\partial P_{ij}}{\partial \delta_i}\right|_{(\delta_i=0)} = K_{4ij} \tag{95}$$

$$\left.\frac{\partial^2 J_k}{\partial \delta_k^{\,2}}\right|_{(\delta_k=0)} = 2 \tag{96}$$

$$K_2 = 2\mu + G_2 \tag{97}$$

$$K_{4ij} = \mu + \rho_{ij} GC_o \tag{98}$$

If it is assumed that ρ_{ij} is independent of δ_i and the variation of the loss due to the variation in ρ is small, then K_{4ij} can be approximated by:

$$K_{4ij} = \mu + \rho\ GC_o \tag{99}$$

where

ρ is the expected value of the correlation coefficients between all consumers i and j

The variation in the loss due to the variation in ρ was found to be small in comparison with the variation of μ_k and is ignored in this calculation procedure. The variation in ρ can be included using the calculation procedure described in section 6.4.

v. Calculate the first and second Taylor components of the variance:

$$T1_k = \sum_{s=1}^{Sections} \left[\begin{array}{l} \left.\dfrac{\partial J_k}{\partial \delta_k}\right|_{(\delta_k=0)} (Qa_{s,k} + Qb_{s,k} + Qc_{s,k})(Rp_s + Rn_s) \\ + 2\sum\limits_{\substack{i=1 \\ i \neq k}}^{N} \left.\dfrac{\partial P_{ij}}{\partial \delta_i}\right|_{(\delta_i=0)} (Qa_{s,k}Qa_{s,i} + Qb_{s,k}Qb_{s,i} + Qc_{s,k}Qc_{s,i})(Rp_s + Rn_s) \\ - \sum\limits_{i=1}^{N} \left.\dfrac{\partial P_{ij}}{\partial \delta_i}\right|_{(\delta_i=0)} (Qa_{s,k}Qb_{s,i} + Qb_{s,k}Qc_{s,i} + Qb_{s,k}Qc_{s,i})Rn_s \\ - \sum\limits_{i=1}^{N} \left.\dfrac{\partial P_{ij}}{\partial \delta_i}\right|_{(\delta_i=0)} (Qa_{s,i}Qb_{s,k} + Qb_{s,i}Qc_{s,k} + Qb_{s,i}Qc_{s,k})Rn_s \end{array} \right] \quad (100)$$

and

$$T2_k = \sum_{s=1}^{Sections} \left.\frac{\partial^2 J_k}{\partial \delta_k{}^2}\right|_{(\delta_k=0)} (Qa_{s,k} + Qb_{s,k} + Qc_{s,k})(Rp_s + Rn_s) \quad (101)$$

vi. Calculate the variance of the average losses on the feeder:

$$V[\mu_{loss}] \approx \sum_{k=1}^{N} T1_k^2 V\delta + \frac{T2_k^2}{4} V\delta 2 + T1_k T2_k E[\delta_k^3] \quad (102)$$

Section 5.6 contains a MathCAD worksheet of this equation with a sample calculation.

5.4 TESTING

The method was tested using a similar process to that described in section 4.6. The methods were tested on a medium income community (Claremont) for the month of July and on a poor community (Tambo) for five months (April – August, inclusive).

5.4.1 Balanced three phase feeder with load data from a low income community

The following table specifies a three phase feeder which was used to test the loss calculation method.

Table 15 A balanced three phase feeder used to test the loss calculation method

	Consumer connections			Conductor resistance per section (Ω)	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1	3			0.0868	0.0868
2		3		0.0868	0.0868
3			3	0.0868	0.0868
4			3	0.0868	0.0868
5		3		0.0868	0.0868
6	3			0.0868	0.0868
7	3			0.0868	0.0868
8		3		0.0868	0.0868
9			3	0.0868	0.0868
10			3	0.0868	0.0868
11		3		0.0868	0.0868
12	3			0.0868	0.0868

The following table compares the calculated and simulated results for the balanced three-phase network using load data from a low-income community (Tambo).

Table 16 Comparison of simulated and calculated results

Parameter	Simulated value [W]	Calculated value [W]	Difference [W]
$E[\mu_{loss}]$	62.2	61.9	0.3
$\sigma_{\mu loss}$	24.1	23.7	0.4

The difference between the calculated and simulated values is small and considered acceptable given the uncertainty in the Monte Carlo simulation results. 1000 Monte Carlo iterations were used to estimate the average and standard deviation of the loss. This results in a error at of 1.25W at 90% confidence.

5.4.2 A single phase feeder with load data from a poor community (Tambo)

The following table specifies a single phase feeder which was used to test the loss calculation method.

Table 17 A single phase network used to test the technical loss calculation method

	Consumer connections			Conductor resistance per section (Ω)	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1	3			0.043	0.043
2	3			0.043	0.043
3	3			0.043	0.043
4	3			0.043	0.043
5	3			0.043	0.043
6	3			0.043	0.043
7	3			0.043	0.043
8	3			0.043	0.043
9	3			0.043	0.043
10	3			0.043	0.043
11	3			0.043	0.043
12	3			0.043	0.043

The following table compares the calculated and simulated results for the single-phase network using load data from a low-income community (Tambo).

Table 18 Comparison of simulated and calculated results

Parameter	Simulated value [W]	Calculated value [W]	Difference [W]
$E[\mu_{loss}]$	67.8	66.2	1.6
$\sigma_{\mu loss}$	32.7	32.0	0.7

The difference between the calculated and simulated values is small and considered acceptable given the uncertainty in the Monte Carlo simulation results. 1000 Monte Carlo iterations were used to estimate the average and standard deviation of the loss. This resulted in a error of 1.7W at 90% confidence.

5.4.3 A short three phase feeder with load data from a medium income community (Claremont)

The following table specifies a short three phase feeder which was used to test the loss calculation method.

Table 19 A short three phase feeder used to test the technical loss calculation method

	Consumer connections			Conductor resistance per section (Ω)	
Node	ma	mb	mc	Rp	Rn
1	1	1	1	0.0868	0.0868
2	1	1	1	0.0868	0.0868
3	1	1	1	0.0868	0.0868

The following table compares the calculated and simulated results for the short three phase feeder using load data from a medium income community (Claremont).

Table 20 Comparison of simulated and calculated results

Parameter	Simulated value [W]	Calculated value [W]	Difference [W]
$E[\mu_{loss}]$	385.6	398.7	-13.1
$\sigma_{\mu loss}$	111.4	113.0	-1.6

The difference between the calculated and simulated values is 3%. The error due to the small number of Monte Carlo iterations (1000) used to estimate the loss is 1.5%. A further source of error in the simulations is that only 72 consumers are available. One of the assumptions of the loss calculation method is that a continuous distribution of individual load current distributions is present and with only 72 consumers the distribution is really discreet (see section 3.10 for a general discussion). This is a problem that is more likely to have an effect on the results when the number of consumers on the feeder is small. In the light of these uncertainties, the difference between the results is accepted for practical design purposes.

5.5 SUMMARY AND CONCLUSIONS

A method for calculating the expected value and the variance of the average technical losses on a residential feeder is described in this section. The method was tested using load data from a low and a medium income community. Simulated results were obtained using a Monte Carlo type process and compared with the calculated results. The calculated and simulated results compare well if one considers the range of uncertainties to which the results are subjected. Some of the uncertainties are:

- Small number of Monte Carlo simulations
- Small number of consumers available from which to draw Monte Carlo samples

The result of the calculations can be used to estimate the losses component of the life cycle cost of a feeder.

Three full examples of the use of the method is given, including input parameters, intermediate results and final results.

5.6 APPENDIX: BALANCED THREE PHASE FEEDER CALCULATIONS

This section contains the input data and sample calculations for the balanced thre phase feeder specified in section 5.4.1.

5.6.1 Input parameters

$V\delta := 0.15$	$V\delta 2 := 0.22$	$E\delta 3 := 0.15$
$G := 1.41$	$C_o := 0.63$	$\rho := 0.07$
$G2 := 3.89$	$C_{o2} := 0.82$	$\mu := 0.322$
$N := 36$	$\text{Sections} := 12$	

5.6.2 Calculated coefficients

$E[K1] := \mu^2 + C_{o2}$ $\qquad$ $E[K1] = 0.924$

$E[K3\_ij] := \mu^2 + \rho \cdot C_o^{\ 2}$ $\qquad$ $E[K3\_ij] = 0.131$

$E[J\_k] := E[K1] + V\delta$ $\qquad$ $E[J\_k] = 1.074$

$E[P\_ij] := E[K3\_ij]$ $\qquad$ $E[P\_ij] = 0.131$

$K2 := 2\mu + G2$ $\qquad$ $K2 = 4.534$

$K4\_ij := \mu + \rho \cdot G C_o$ $\qquad$ $K4\_ij = 0.384$

$D[J\_k(0)] := K2$

$D[P\_ij(0)] := K4\_ij$

$D2[J\_k(0)] := 2$

5.6.3 Calculation of expected value of average loss

$$E[Loss1] := \sum_{s=0}^{Sections-1}\left[\sum_{k=0}^{N-1} E[J\_k]\left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}}\right)\cdot\left(Rp_s + Rn_s\right)\right]$$

$E[Loss1] = 43.616$

$$E[Loss2] := \sum_{s=0}^{Sections-1}\left[\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i \neq j)\cdot\left[E[P\_ij]\left(Q_{a_{i,s}}\cdot Q_{a_{j,s}} + Q_{b_{i,s}}\cdot Q_{b_{j,s}} + Q_{c_{i,s}}\cdot Q_{c_{j,s}}\right)\cdot\left(Rn_s + Rp_s\right)\right]\right]$$

$E[Loss2] = 40.259$

$$E[Loss3] := \sum_{s=0}^{Sections-1}\left[-\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i \neq j)\cdot\left[E[P\_ij]\left(Q_{a_{i,s}}\cdot Q_{b_{j,s}} + Q_{a_{i,s}}\cdot Q_{c_{j,s}} + Q_{b_{i,s}}\cdot Q_{c_{j,s}}\right)\cdot Rn_s\right]\right]$$

$E[Loss3] = -21.978$

5.6.4 Calculation of variance of average loss

$$T_{1a}(k) := \sum_{s=0}^{Sections-1}\left[D[J\_k(0)]\left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}}\right)\cdot\left(Rp_s + Rn_s\right)\right]$$

$$T_{1b}(k) := \sum_{s=0}^{Sections-1} 2\left[\sum_{i=0}^{N-1}(i \neq k)\cdot D[P\_ij(0)]\left(Q_{a_{k,s}}\cdot Q_{a_{i,s}} + Q_{b_{k,s}}\cdot Q_{b_{i,s}} + Q_{c_{k,s}}\cdot Q_{c_{i,s}}\right)\cdot\left(Rp_s + Rn_s\right)\right]$$

$$T_{1c}(k) := \sum_{s=0}^{Sections-1} -\left[\sum_{i=0}^{N-1} D[P\_ij(0)]\left(Q_{a_{k,s}}\cdot Q_{b_{i,s}} + Q_{b_{k,s}}\cdot Q_{c_{i,s}} + Q_{a_{k,s}}\cdot Q_{c_{i,s}} + Q_{a_{i,s}}\cdot Q_{b_{k,s}} + Q_{b_{i,s}}\cdot Q_{c_{k,s}} + Q_{a_{i,s}}\cdot Q_{c_{k,s}}\right)\cdot\left(Rn_s\right)\right]$$

$$T_1(k) := T_{1a}(k) + T_{1b}(k) + T_{1c}(k)$$

$$T_2(k) := \sum_{s=0}^{Sections-1}\left[D2[J\_k(0)]\left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}}\right)\cdot\left(Rp_s + Rn_s\right)\right]$$

5.6.5 Calculated $T_1(k)$ and $T_2(k)$ for $k = 0$ to 35

k	T1(k)	T2(k)	k	T1(k)	T2(k)	k	T1(k)	T2(k)
0	13.05	4.17	12	9.93	2.78	24	5.92	1.39
1	13.05	4.17	13	9.93	2.78	25	5.92	1.39
2	13.05	4.17	14	9.93	2.78	26	5.92	1.39
3	12.49	3.82	15	8.78	2.43	27	4.66	1.04
4	12.49	3.82	16	8.78	2.43	28	4.66	1.04
5	12.49	3.82	17	8.78	2.43	29	4.66	1.04
6	12.04	3.47	18	7.72	2.08	30	3.01	0.69
7	12.04	3.47	19	7.72	2.08	31	3.01	0.69
8	12.04	3.47	20	7.72	2.08	32	3.01	0.69
9	11.19	3.13	21	6.77	1.74	33	1.45	0.35
10	11.19	3.13	22	6.77	1.74	34	1.45	0.35
11	11.19	3.13	23	6.77	1.74	35	1.45	0.35

5.6.6 Results

$$E[\mathrm{Loss}] := E[\mathrm{Loss1}] + E[\mathrm{Loss2}] + E[\mathrm{Loss3}] \qquad E[\mathrm{Loss}] = 61.482$$

$$V[\mathrm{Loss}] := \sum_{i=0}^{N-1} \left(V\delta \cdot T_1(i)^2 + V\delta 2 \frac{T_2(i)^2}{4} + E\delta 3 \cdot T_1(i) \cdot T_2(i) \right) \qquad \sqrt{V[\mathrm{Loss}]} = 23.625$$

5.6.7 Qa

k	1	2	3	4	5	6	7	8	9	10	11	12
0	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	1	1	1	0	0	0	0	0
16	1	1	1	1	1	1	1	0	0	0	0	0
17	1	1	1	1	1	1	1	0	0	0	0	0
18	1	1	1	1	1	1	0	0	0	0	0	0
19	1	1	1	1	1	1	0	0	0	0	0	0
20	1	1	1	1	1	1	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0
33	1	0	0	0	0	0	0	0	0	0	0	0
34	1	0	0	0	0	0	0	0	0	0	0	0
35	1	0	0	0	0	0	0	0	0	0	0	0

5.6.8 Qb

k	1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1	1	0
4	1	1	1	1	1	1	1	1	1	1	1	0
5	1	1	1	1	1	1	1	1	1	1	1	0
6	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	0	0	0	0
13	1	1	1	1	1	1	1	1	0	0	0	0
14	1	1	1	1	1	1	1	1	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0
21	1	1	1	1	1	0	0	0	0	0	0	0
22	1	1	1	1	1	0	0	0	0	0	0	0
23	1	1	1	1	1	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0
30	1	1	0	0	0	0	0	0	0	0	0	0
31	1	1	0	0	0	0	0	0	0	0	0	0
32	1	1	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0

5.6.9 Qc

k	1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1	0	0
7	1	1	1	1	1	1	1	1	1	1	0	0
8	1	1	1	1	1	1	1	1	1	1	0	0
9	1	1	1	1	1	1	1	1	1	0	0	0
10	1	1	1	1	1	1	1	1	1	0	0	0
11	1	1	1	1	1	1	1	1	1	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0
24	1	1	1	1	0	0	0	0	0	0	0	0
25	1	1	1	1	0	0	0	0	0	0	0	0
26	1	1	1	1	0	0	0	0	0	0	0	0
27	1	1	1	0	0	0	0	0	0	0	0	0
28	1	1	1	0	0	0	0	0	0	0	0	0
29	1	1	1	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0

5.7 APPENDIX: SINGLE PHASE FEEDER CALCULATIONS

The calculations for the single phase feeder is given in this section. The values for Qa, Qb and Qc is not shown.

5.7.1 Input parameters

$N := 36$ $\quad$ $Sections := 12$

$\mu := 0.28$ $\quad$ $\rho := 0.08$

$V\delta := 0.15$

$G := 1.54$ $\quad$ $C_o := 0.5$

$G2 := 4.55$ $\quad$ $C_{o2} := 0.7$

$V\delta2 := 0.32$ $\quad$ $E\delta3 := 0.179$

5.7.2 Initial calculations

$E[K1] := \mu^2 + C_{o2}$ $\quad$ $E[K1] = 0.778$

$E[K3\_ij] := \mu^2 + \rho \cdot C_o{}^2$ $\quad$ $E[K3\_ij] = 0.098$

$E[J\_k] := E[K1] + V\delta$ $\quad$ $E[J\_k] = 0.928$

$E[P\_ij] := E[K3\_ij]$ $\quad$ $E[P\_ij] = 0.098$

$K2 := 2\mu + G2$ $\quad$ $K2 = 5.11$

$K4\_ij := \mu + \rho \cdot G \cdot C_o$ $\quad$ $K4\_ij = 0.342$

$D[J\_k(0)] := K2$

$D[P\_ij(0)] := K4\_ij$

$D2[J\_k(0)] := 2$

5.7.3 Calculations for expected values

$$E[Loss1] := \sum_{s=0}^{Sections-1} \left[\sum_{k=0}^{N-1} E[J\_k] \left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$E[Loss1] = 18.683$

$$E[Loss2] := \sum_{s=0}^{Sections-1} \left[\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Q_{a_{i,s}} \cdot Q_{a_{j,s}} + Q_{b_{i,s}} \cdot Q_{b_{j,s}} + Q_{c_{i,s}} \cdot Q_{c_{j,s}} \right) \cdot \left(Rn_s + Rp_s \right) \right] \right]$$

$E[Loss2] = 47.525$

$$E[Loss3] := \sum_{s=0}^{Sections-1} \left[-\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Q_{a_{i,s}} \cdot Q_{b_{j,s}} + Q_{a_{i,s}} \cdot Q_{c_{j,s}} + Q_{b_{i,s}} \cdot Q_{c_{j,s}} \right) \cdot Rn_s \right] \right]$$

$E[Loss3] = 0$

5.7.4 Calculation of Taylor terms

$$E[Loss1] := \sum_{s=0}^{Sections-1} \left[\sum_{k=0}^{N-1} E[J\_k] \left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$E[Loss1] = 18.683$

$$E[Loss2] := \sum_{s=0}^{Sections-1} \left[\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Q_{a_{i,s}} \cdot Q_{a_{j,s}} + Q_{b_{i,s}} \cdot Q_{b_{j,s}} + Q_{c_{i,s}} \cdot Q_{c_{j,s}} \right) \cdot \left(Rn_s + Rp_s \right) \right] \right]$$

$E[Loss2] = 47.525$

$$E[Loss3] := \sum_{s=0}^{Sections-1} \left[-\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Q_{a_{i,s}} \cdot Q_{b_{j,s}} + Q_{a_{i,s}} \cdot Q_{c_{j,s}} + Q_{b_{i,s}} \cdot Q_{c_{j,s}} \right) \cdot Rn_s \right] \right]$$

$E[Loss3] = 0$

5.7.5 Calculated Taylor terms

k	T1(k)	T2(k)	k	T1(k)	T2(k)	k	T1(k)	T2(k)
0	18.32	2.06	12	15.03	1.38	24	8.93	0.69
1	18.32	2.06	13	15.03	1.38	25	8.93	0.69
2	18.32	2.06	14	15.03	1.38	26	8.93	0.69
3	17.76	1.89	15	13.77	1.20	27	6.96	0.52
4	17.76	1.89	16	13.77	1.20	28	6.96	0.52
5	17.76	1.89	17	13.77	1.20	29	6.96	0.52
6	17.03	1.72	18	12.33	1.03	30	4.82	0.34
7	17.03	1.72	19	12.33	1.03	31	4.82	0.34
8	17.03	1.72	20	12.33	1.03	32	4.82	0.34
9	16.12	1.55	21	10.72	0.86	33	2.50	0.17
10	16.12	1.55	22	10.72	0.86	34	2.50	0.17
11	16.12	1.55	23	10.72	0.86	35	2.50	0.17

5.7.6 Results

$$E[Loss] := E[Loss1] + E[Loss2] + E[Loss3]$$

$$E[Loss] = 66.208$$

$$V[Loss] := \sum_{i=0}^{N-1} \left(V\delta \cdot T_1(i)^2 + V\delta 2 \frac{T_2(i)^2}{4} + E\delta 3 \cdot T_1(i) \cdot T_2(i) \right)$$

$$\sqrt{V[Loss]} = 32.07$$

5.8 APPENDIX: SHORT BALANCED THREE-PHASE FEEDER CALCULATIONS

This section contains the calculations for the short balanced three-phase feeder with load data from a medium income community (Claremont).

5.8.1 Input parameters

$N := 9 \quad Sections := 3$

$V\delta := 13.77 \qquad V\delta 2 := 672.1 \qquad E\delta 3 := 60$

$G2 := 5.81 \qquad C_{o2} := 50.56 \qquad \rho := 0.13$

$G := 0.39 \qquad C_o := 6.887 \qquad \mu := 7.28$

5.8.2 Initial calculations

$E[K1] := \mu^2 + C_{o2}$ $E[K1] = 103.558$

$E[K3\_ij] := \mu^2 + \rho \cdot C_o^{\,2}$ $E[K3\_ij] = 59.164$

$E[J\_k] := E[K1] + V\delta$ $E[J\_k] = 117.328$

$E[P\_ij] := E[K3\_ij]$ $E[P\_ij] = 59.164$

$K2 := 2\mu + G2$ $K2 = 20.37$

$K4\_ij := \mu + \rho \cdot G C_o$ $K4\_ij = 7.629$

$D[J\_k(0)] := K2$

$D[P\_ij(0)] := K4\_ij$

$D2[J\_k(0)] := 2$

5.8.3 Expected value calculations

$$E[Loss1] := \sum_{s=0}^{Sections-1} \left[\sum_{k=0}^{N-1} E[J\_k] \left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$E[Loss1] = 366.628$

$$E[Loss2] := \sum_{s=0}^{Sections-1} \left[\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Q_{a_{i,s}} \cdot Q_{a_{j,s}} + Q_{b_{i,s}} \cdot Q_{b_{j,s}} + Q_{c_{i,s}} \cdot Q_{c_{j,s}} \right) \cdot \left(Rn_s + Rp_s \right) \right] \right]$$

$E[Loss2] = 246.503$

$$E[Loss3] := \sum_{s=0}^{Sections-1} \left[-\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Q_{a_{i,s}} \cdot Q_{b_{j,s}} + Q_{a_{i,s}} \cdot Q_{c_{j,s}} + Q_{b_{i,s}} \cdot Q_{c_{j,s}} \right) \cdot Rn_s \right] \right]$$

$E[Loss3] = -215.69$

5.8.4 Taylor term calculations

$$T_{1a}(k) := \sum_{s=0}^{\text{Sections}-1} \left[D[J\_k(0)] \left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$$T_{1b}(k) := \sum_{s=0}^{\text{Sections}-1} 2 \left[\sum_{i=0}^{N-1} (i \neq k) \cdot D[P\_ij(0)] \left(Q_{a_{k,s}} \cdot Q_{a_{i,s}} + Q_{b_{k,s}} \cdot Q_{b_{i,s}} + Q_{c_{k,s}} \cdot Q_{c_{i,s}} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$$T_{1c}(k) := \sum_{s=0}^{\text{Sections}-1} - \left[\sum_{i=0}^{N-1} D[P\_ij(0)] \left(Q_{a_{k,s}} \cdot Q_{b_{i,s}} + Q_{b_{k,s}} \cdot Q_{c_{i,s}} + Q_{a_{k,s}} \cdot Q_{c_{i,s}} + Q_{a_{i,s}} \cdot Q_{b_{k,s}} + Q_{b_{i,s}} \cdot Q_{c_{k,s}} + Q_{a_{i,s}} \cdot Q_{c_{k,s}} \right) \cdot \left(Rn_s \right) \right]$$

$$T_1(k) := T_{1a}(k) + T_{1b}(k) + T_{1c}(k)$$

$$T_2(k) := \sum_{s=0}^{\text{Sections}-1} \left[D2[J\_k(0)] \left(Q_{a_{k,s}} + Q_{b_{k,s}} + Q_{c_{k,s}} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

5.8.5 Taylor terms results

k	T1(k)	T2(k)
0	10.61	1.04
1	8.40	0.69
2	4.86	0.35
3	10.61	1.04
4	8.40	0.69
5	4.86	0.35
6	10.61	1.04
7	8.40	0.69
8	4.86	0.35

5.8.6 Results

$$E[Loss] := E[Loss1] + E[Loss2] + E[Loss3] \qquad E[Loss] = 397.441$$

$$V[Loss] := \sum_{i=0}^{N-1} \left(V\delta \cdot T_1(i)^2 + V\delta 2 \frac{T_2(i)^2}{4} + E\delta 3 \cdot T_1(i) \cdot T_2(i) \right) \qquad \sqrt{V[Loss]} = 112.831$$

6. FUZZY PROBABILISTIC CALCULATION METHODS

6.1 INTRODUCTION

Little is known about some of the uncertainties in low voltage feeder design. If a minimum, maximum and most likely values are known, the uncertainties can be modelled with fuzzy numbers. The probabilistic calculation methods described in chapters 3,4 and 5 can be modified to include these uncertainties. The results of the probabilistic methods are parameters of a probability distribution. Since fuzzy numbers are essentially ranges of possible values, the result of the fuzzified probabilistic methods are ranges of probability distributions. These ranges of probability distributions may be termed fuzzy probability distributions [20]. A design percentile from a fuzzy probability distribution is therefore not a single value but a range of possible values and can be represented with a fuzzy number.

In this chapter the modifications to the methods described in chapters 4 and 5 are shown. Specifically the following uncertainties are considered

- Phase and neutral conductor resistance variation due to temperature and length specifications
- Supply voltage variation
- Uncertainty about forecasted load parameters

6.2 THEORETICAL DERIVATION

.

Triangular fuzzy numbers were chosen to represent the uncertainties because

- They are mathematically very simple
- Not much more information than a possible range is available for each uncertainty
- The uncertainty about a parameter should increase depending on the level of presumption

This means that the uncertainties can readily be modelled without much information. Note that the use of triangular fuzzy numbers assumes a linear progression from the minimum to the most likely value and also a linear decreasing progression from the most likely to the maximum value. The impact of this assumption could be the topic of further research.

6.2.1 Parameter uncertainty

Triangular fuzzy numbers can be used to describe the uncertainty about the load parameters. Each parameter has the following:

- An absolute maximum
- A likely value
- An absolute minimum

The parameters for the probabilistic calculation methods are

- $\mu_{average}$, the average of the load current means
- $V\delta$, the variance of the load current means
- $V\delta 2$, the variance δ^2
- $E\delta 3$, the expected value of δ^3, or the third moment of δ^3
- $E(\sigma)$, the average of the individual load current standard deviations, this value is equal to C_o
- $E(\sigma^2)$, the average of the individual load current variances, this value is equal to C_{o2}
- $E(\rho)$, the expected value of the individual correlations
- G_2, the slope of the regression line which relates δ_k to σ^2_k
- G, the slope of the regression line which relates δ_k to σ_k

Note that some of the load parameters are uncertain due to errors in the underlying models and in the specification. These errors can be combined into the total parameter uncertainty.

6.2.2 Neutral and phase resistance uncertainty

The resistance of the neutral and phase conductors of an LV feeder might not be exactly as planned. This could be due to physical discrepancy between actual and design length. The operating temperature may also differ from the planned operating temperature. A triangular fuzzy number can be used to represent the uncertainty about the conductor resistance[19].

6.2.3 Supply voltage uncertainty

The value of the supply voltage is also uncertain and may vary from the assumed value. The supply voltage can be represented with a triangular fuzzy number with minimum, maximum and most likely value.

6.3 FUZZY PROBABILISTIC VOLTAGE REGULATION

Chapter 4 describes a probabilistic calculation method to calculate the voltage performance on an LV feeder with residential customers. The result of the procedure is a percentile of a voltage performance distribution at some level of confidence. If the level of confidence is predefined, say 90%, then the minimum and maximum values of the range can be calculated. The most likely value is calculated using the design values for the load parameters, neutral resistance, phase resistance and supply voltage.

One common way of specifying the fuzzy uncertainties is percentage movement from the design value (e.g. ±5%).

The likely value of the voltage performance design percentile is obtained using the likely values of all the fuzzy parameters. The voltage performance calculations have a non-linear, non-monotonic behaviour when considering the marginal contribution of each uncertainty. This requires the use of either a non-linear programming approach or a solution search approach.

6.3.1 Non-linear programming solution

The range of voltage performance percentile values due to the fuzzy uncertainties can be solved using a non-linear programming solution. Several algorithms for non-linear programming are available[40] with different properties, convergence times etc.

The non-linear programming method chosen to evaluate the minimum and maximum design percentiles was the conjugate gradient method. This method is considered among the best general purpose methods available [40 p238]. This method is implemented in MathCAD 2000 is able to solve the problem in a reasonable time (see the testing section for typical times).

The choice of an optimal solution method could be the topic of further research.

The fuzzy probabilistic result can be obtained using the following steps:

1. Obtain a function, $V_{\%tile}$, which uses a set of parameters, p, and returns the voltage performance design percentile for some feeder
2. For each parameter the upper and lower value of the fuzzy number is defined as constraints to the non-linear programming problem
3. Minimize and Maximize $V_{\%tile}$ given the set of constraints
4. The most likely value for $V_{\%tile}$ is calculated using the likely values from the set of parameters.

6.3.2 Exhaustive and genetic algorithm

The range of voltage performance percentile values may be calculated using an exhaustive search using all the possible input values. If it is assumed that the marginal effect of each uncertainty is monotonic, then the minimum and maximum values of each of the ranges of possible values could be tried until all the possible combination of minimums and maximums are exhausted. The minimum and maximum of the design percentiles can then be extracted from the range. A total of 2^N combinations of input parameters need to be tested, where N is the total number of uncertain variables. The assumption in this approach could lead to sub optimal solutions to the problem. However, this problem is not likely to occur on the heaviest loaded phase of the feeder.

An alternative approach is to use a genetic algorithm to search for the minimum and maximum values of the design percentiles [47].

Exhaustive search

To solve for the minimum and maximum of the design percentiles, the following steps can be followed:

1. Define a function, $V_{\%tile}$, dependent on a set of parameters, p. The function returns a design percentile given p
2. Assign a new set of values to p

3. Evaluate $V_{\%tile}$ and compare the result with previous calculations. Retain the minimum and maximum values
4. Repeat 2 and 3, until all possible combination of parameters have been tried.

The solution is obtained after all possible values for p has been applied and the time to completion is .

Genetic algorithm (GA)

The work on genetic algorithms is intended to be exploratory of nature and the aim is to illustrate that the solution can be found through genetic algorithms. This section is not intended as a comprehensive investigation into the application of GA to the solution of this type of problem. More work is required before GA's can be robustly applied to the problem. This section does however identify the potential of GA's to provide a practical solution.

In genetic algorithms, possible solutions are encoded as chromosomes. An initial set of problems are either randomly generated or can be specified. This initial set is called the population. The search for a solution progress in steps and each step is called a generation, since a new set of possible solutions is generated through the following mechanism:

- Part of the solution is randomly altered – mutation
- Two solutions are combined – recombination

The probability of mutation in a particular part of the chromosome needs to be specified. The probability of recombination and the mechanism for recombination also need to be specified.

Each member of a generation is evaluated to obtain a number representing its fitness. In the voltage percentile problem, this would be the relative rating of the voltage percentile. For instance, if the maximum voltage percentile is sought, then the largest voltage percentile in a generation has the largest fitness, the second largest has the second largest fitness etc. A portion of the chromosomes is discarded based on fitness.

A very general description of a genetic algorithm is as follows [15 p5]:

1. Initialise a population of chromosomes
2. Evaluate each chromosome in the population
3. Create new chromosomes by mating current chromosomes; apply mutation and recombination as the parent chromosomes mate
4. Delete members of the population to make room for the new chromosomes
5. Evaluate the new chromosomes and insert them into the population
6. If time is up, stop and return the best chromosome; if not, go to 3

After a predefined period of time, the algorithm is stopped and the best solution is taken to be the final solution.

The genetic algorithm out performs the exhaustive search in terms of finding a good solution. Although the genetic algorithm does not always find the absolute minimum or maximum values, it does return values close to them. Specification of the mutation probability and recombination mechanism influences the performance of the method.

An advantage of the genetic algorithm is that it is faster than non-linear programming for a large number of uncertainties (increased complexity) and exhaustive search algorithms.

6.4 FUZZY PROBABILISTIC LOSS CALCULATION METHOD

The minimum and maximum values of the range of design percentiles of the loss calculation can easily be found since the set of equations in the loss calculation method is additive. It should be noted that the additive nature is due to the underlying assumptions about the load behaviour (constant current etc). If any of these assumptions are violated then this method would not be valid and an approach similar to the one described for the fuzzy-probabilistic voltage performance method should be used. To derive the minimum design percentile it is simply necessary to calculate the design percentile using all the minimum parameter values. Using a similar approach, the maximum design percentile can be obtained.

The process of calculation can be summarized as:

1. Determine the minimum value of all the fuzzy uncertainties
2. Calculate the loss design percentile using the set of minimum values
3. Determine the maximum value of all the fuzzy uncertainties
4. Calculate the loss design percentile using the set of maximum values

The fuzzy probabilistic result is therefore

$$L_{\%tile}=(Lfunc[Min(P)];Lfunc[Likely(P)];Lfunc[Max(P)]) \qquad (103)$$

Where

$L_{\%tile}$ is a fuzzy number representing the design percentiles at some level of confidence

Lfunc is a function which returns the design loss percentile value given a set of parameters p

P is a set of fuzzy parameters, each with a minimum, maximum and likely value

Min is a function returning only the minimum values of each of the fuzzy parameters

Max is a function returning only the maximum values of each of the fuzzy parameters

Likely is a function returning only the likely values of each of the fuzzy parameters

6.5 TESTING

Figure 23 contains an illustration of a short balanced feeder used to test the fuzzy probabilistic methods. Load parameters from a medium income community (Claremont) were used in the calculations.

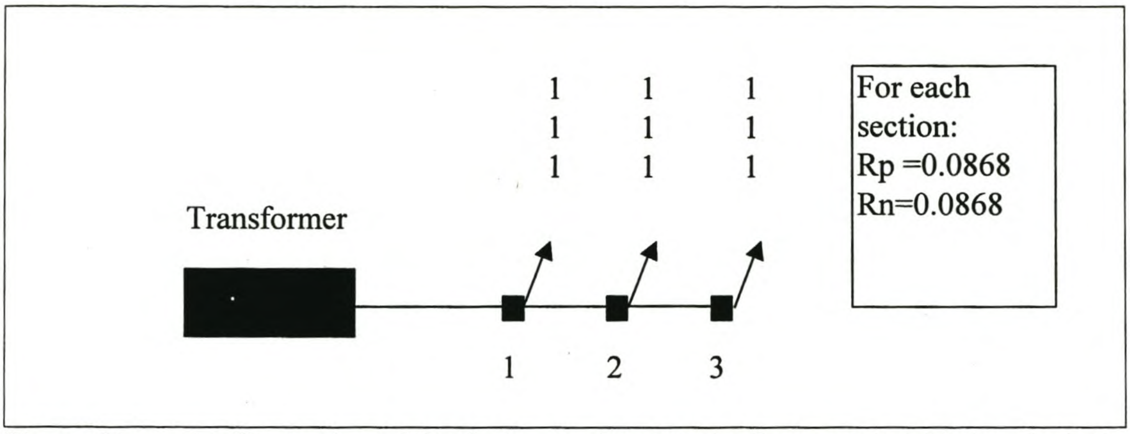

Figure 23 An illustration of the feeder used to test the fuzzy-probabilistic calculation method. A description of the feeder is given in the table below.

Table 21 A short three phase network used to test the fuzzy probabilistic voltage performance calculation method

	Number of consumers			Conductor resistance per section (Ω)	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1	1	1	1	0.0868	0.0868
2	1	1	1	0.0868	0.0868
3	1	1	1	0.0868	0.0868

For the test, an uncertainty of ±10% on all the load parameters, the supply voltage and the conductor resistance was assumed.

6.5.1 Fuzzy probabilistic voltage performance calculation

A fuzzy probabilistic voltage performance method results in a fuzzy number with a minimum, a maximum and a most likely value. Table 22 shows a comparison of results that were calculated using the following methods:

a) non-linear programming
b) genetic algorithm

Table 22 Comparison of the results of the fuzzy probabilistic voltage calculation method using non-linear programming and a genetic algorithm (All results in V)

Results	Non-linear programming	Genetic algorithm		Exhaustive search
		5 generations	10 generations	
Average of Minimum	179.4	179.6	179.5	179.4
Standard deviation of error	N/A	0.27	0.23	N/A
Average of Maximum	233.4	233.1	233.2	233.4
Standard deviation of error	N/A	0.28	0.24	N/A
Relative time to results	1	0.2	0.4	0.5

The genetic algorithm was stopped after 5 and 10 generations. In both cases, the standard deviation of the error for the minimum and maximum values were calculated and is shown in the table. The genetic algorithm performed well with the following parameters:
Population size = 8; Mutation rate = 0.4 ; Cross over points = 1

A comparison of the relative time to results indicates that genetic algorithms could be used to get an estimate of the upper and lower value of the range of design percentiles within a short time. The choice of parameters for the genetic algorithm influences the performance. Future research may be conducted to obtain an optimum set of parameters.

Accurate results can be calculated using a non-linear programming approach or an exhaustive search. An advantage of an exhaustive search is that it can calculate the minimum and maximum values at the same time.

A software implementation of the genetic algorithm by [27] was used to obtain the results. The non-linear programming was solved using the conjugate gradient method as implemented in MathCAD 2000.

6.5.2 Fuzzy probabilistic loss calculation

The results of the fuzzy probabilistic loss calculations are obtained through three separate calculations using the method described in chapter 5. The first calculation is with the minimum of each of the load parameters and the conductor resistances. The result of the first calculation is the minimum value of the calculated fuzzy number. Similarly the most likely and maximum results can be found.

Table 23 compares results found using the method described above with the results from an exhaustive search of the problem space.

Table 23 Comparison of the results of the fuzzy probabilistic loss calculation method and a exhaustive search of the problem space

Results	Exhaustive search	Proposed method
Minimum E[Loss] in W	305.8	305.8
Maximum E[Loss] in W	505.1	505.1
Minimum V[Loss] in W^2	7 845.7	7 845.7
Maximum V[Loss] in W^2	19 802.1	19 802.1

The calculation in three steps (minimum, most likely, maximum) is very fast and exact results are obtained.

6.6 SUMMARY AND CONCLUSIONS

Some of the uncertainties in residential feeder design can best be represented with fuzzy numbers. A method for calculating the effect of these uncertainties on the results of the probabilistic methods, is described and some tests results are shown. Combining these concepts results in a fuzzy probabilistic approach.

The fuzzy probabilistic voltage performance can be calculated using any one of the following methods:

- Non-linear programming
- Genetic algorithms
- Exhaustive search of the problem space

The non-linear programming and exhaustive search methods give exact results, but are time consuming. The genetic algorithms can quickly return results, but the results may not be that accurate. Depending on the complexity of the non-linear equations, it would be possible for the non-linear programming solution to out-perform both the genetic algorithms and exhaustive search approaches. The introduction of the genetic algorithms as a solution was very exploratory and aimed at investigating its usefulness. This investigation is in no way complete and more work could be done to improve the solution finding efficiency for this particular problem.

A method for calculating the fuzzy probabilistic loss is proposed and was compared with the results of an exhaustive search of the problem space. The proposed method is analytical, fast and gives exact results.

7. LOAD PARAMETER SPECIFICATION

7.1 INTRODUCTION

In this chapter the relationship between the load parameters for the voltage drop calculation method and loss calculation method is described. The relationship is investigated through linear regression analysis. An estimate of a complete set of load parameters can be obtained, from the result of the regression analysis, using

- an estimate of the average consumption of a community
- the variation between the consumption of individual consumers.

The identification of the relationships between parameters simplifies the specification of the load parameters and reduces the load parameter uncertainty.

All of the analyses were performed using data collected from 1994 to 1999 by the residential load research projects in South Africa [53, 61] (see section 1.4 for more detail). Load data from 28 communities, with a variety of average household income and average years with electrification, were used in the analysis. A table containing the load parameters used in the regression analyses is included in section 7.7. Some indication of income level and time with electricity is also shown in the table.

All of the lines were fitted using linear regression. In some cases the fit might be improved using non-linear or local regression techniques, this was not investigated and might be a topic of future research. An option that was not explored is the use of neural networks for the prediction or identification of the prediction models. Most of the relationships have a very strong linear component judging by the R^2 values achieved and linear regression generally provides simpler models. Some additional work can be done to improve the robustness of the fitted models.

7.2 CLASSSIFICATION OF LOAD PARAMETERS

The parameters required for the voltage drop calculation method and the loss calculation method can be divided into three categories as follows:

- Scale
- Shape
- Correlation

Two of the parameters has a shape and a correlation component, these are

- G_2, the slope of the regression line which relates δ_k to $\sigma^2{}_k$
- G, the slope of the regression line which relates δ_k to σ_k

The slopes of the regression lines can be rewritten as:

$$G_2 = \rho_{\delta\sigma^2}\sqrt{\frac{V\sigma^2}{V\delta}} \tag{104}$$

where

$V\sigma^2$	is the variance of the individual consumer load current variances
$\rho_{\delta\sigma^2}$	is the correlation between the individual consumer load current averages and variances
$V\delta$	is the variance of the individual consumer load current averages

Similarly G, can be written as

$$G = \rho_{\delta\sigma}\sqrt{\frac{V\sigma}{V\delta}} \tag{105}$$

where

$V\sigma$	is the variance of the individual consumer load current standard deviations
$\rho_{\delta\sigma}$	is the correlation between the individual consumer load current averages and standard deviations
$V\delta$	is the variance of the individual consumer load current averages

Elements of the three categories of parameters are described as:

Scale parameters

- $\mu_{average}$, the average of the load current means, this is an ampere equivalent for the average energy of the consumers in a community
- $E[\sigma]$, the average of the individual load current trace standard deviations
- $E[\sigma^2]$, the average of the individual load current trace variances

Shape parameters

The shape parameters are per-unit variances, second- and third moments.

- CV_{μ}, the coefficient of variation of the load current means
- $V\delta 2/\mu_{average}$, the ratio between the variance of δ^2 and $\mu_{average}$
- $E\delta 3/\mu_{average}$, the ratio between the expected value of δ^3 and $\mu_{average}$
- CV_{σ}, the coefficient of variation of the load current standard deviations
- CV_{σ^2}, the coefficient of variation of the load current variance

Correlation parameters

- $E[\rho]$, the expected value of the individual correlations
- $\rho_{\delta\sigma}$, the correlation between the individual load current means and load current standard deviations
- $\rho_{\delta\sigma^2}$, the correlation between the individual load current means and load current variances

7.3 SCALE PARAMETERS

7.3.1 Specification of $\mu_{average}$

A prediction model for $\mu_{average}$ of consumers in South Africa has been developed by [17]. This model has been implemented in software and is known as the DT Pre Electrification Tool. An inflation-adjusted estimate of monthly household income is used as input to the model for estimating future values of $\mu_{average.}$ A predication horizon of up to 15 years can be specified in the model. The model is based on a non-linear regression which explains more the 95% of the variation in the source data.

A model relating $\mu_{average}$ with the after diversity maximum demand (ADMD) was also developed by [17]. The relationship between the ADMD and $\mu_{average}$ is influenced by the climate of the region the community is located. A hand-scored climatic severity index is used to introduce the difference in climates in found in South Africa. The climatic severity index or CSI varies between –1 and 1, with 1 being very severe (winter rain fall) and –1 being moderate (low temperature variation, high humidity, high average temperature). The model is based on a regression line which explains more than 96% of the variation in the source data.

7.3.2 Specification of E[σ]

The following graph shows the relationship between $\mu_{average}$ and E[σ].

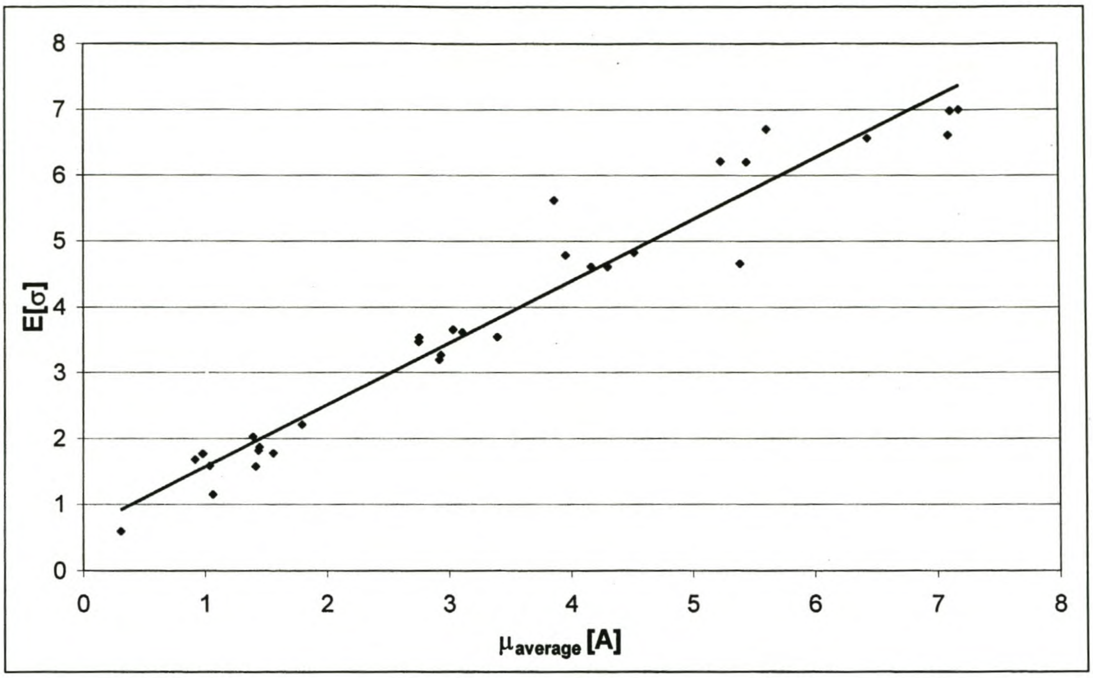

Figure 24 The relationship between $\mu_{average}$ and E[σ]

The relationship, shown in the above figure, is influenced by the climatic severity index (CSI), giving the following equation for the regression line between E[σ], $\mu_{average}$ and CSI:

$$E[\sigma] = 0.91\mu_{average} + 0.25CSI + 0.78 \tag{106}$$

The regression explains 96 % of the variation in the source data and has a standard error of 0.4. The relationship is very strong and should provide good estimates of E[σ].

7.3.3 Specification of E[σ²]

An equally strong relationship was found between E[σ²], CSI and $\mu_{average}$. The equation for the regression line between E[σ²], $\mu_{average}$ and CSI is

$$E[\sigma^2] = 7.89\mu_{average} + 2.64CSI - 6.09 \tag{107}$$

The regression explains 96 % of the variation in the source data and has a standard error of 3.37. The relationship is very strong and should provide good estimates of $E[\sigma^2]$.

7.3.4 Specification of $E[\delta^3]$ and $V[\delta^2]$

A beta distribution can be fitted to the distribution of the sampled load current means. A Kolmogorov-Smirnov test was performed and the theoretical distribution could be accepted. The third and fourth moments of the fitted beta distribution can be obtained using the following equations:

$$E[\delta^3] = \frac{\alpha(\alpha+1)(\alpha+2)}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+2)} C_k^3 \tag{108}$$

$$E[\delta^4] = \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{(\alpha+\beta)(\alpha+\beta+1)(\alpha+\beta+2)(\alpha+\beta+3)} C_k^4 \tag{109}$$

where

α,β are the parameters of the beta distribution and can be calculated using the following equations:

$$\mu_{ns} = \frac{\mu_{average} - \text{Minimum}}{\text{Maximum} - \text{Minimum}} \tag{110}$$

$$\sigma_{ns} = \frac{\sqrt{V\delta}}{\text{Maximum} - \text{mininimum}} \tag{111}$$

Minimum = 0 and Maximum = C_k the circuit breaker size

$$\alpha = \frac{\mu_{ns}{}^2 - \mu_{ns}(\sigma_{ns}^2 + \mu_{ns}^2)}{\sigma_{ns}^2} \tag{112}$$

$$\beta = \frac{\alpha(1 - \mu_{ns})}{\mu_{ns}} \tag{113}$$

$V[\delta]$ can then be calculated using the following equation:

$$V[\delta^2] = E[\delta^4] - (E[\delta^2])^2 \tag{114}$$

7.4 SHAPE PARAMETERS

7.4.1 Specification of CV_μ

Variation of the load current means within a community is very difficult to predict from other measurable community characteristics. A linear regression on the coefficient of variation of the floor area of houses in different communities and the coefficient of variation of the consumption, explains 73% of the variation in the source data. The regression has a standard error of 0.17.

This regression line has the following equation:

$$CV_\mu = 1.1CV_{FloorArea} - 0.07\mu_{average} + 0.52 \tag{115}$$

The relationship is not very strong and could lead to inaccurate estimates of CV_μ. Further, this parameter is influenced by factors, which either have not been measured in the communities used in the modelling or are difficult to measure.

In practice recommended values should be derived for communities which could be classified as similar, different and very different. The choice of CV_μ for each category should be the topic of future research.

7.4.2 Specification of CV_σ

A strong relationship has been found between CV_μ and CV_σ. A linear regression on CV_μ and CV_σ explains 93% of the variation in the source data. The relationship is influenced by the climatic severity index (CSI). The regression has a standard error of 0.06 .

The regression line has the following equation:

$$CV_\sigma = 0.67CV_\mu - 0.03CSI \tag{116}$$

This relationship is strong and with a knowledge of CV_μ should give good estimates of CV_σ.

7.4.3 Specification of $CV\sigma^2$

A strong relationship was found between CV_μ and $\sqrt{(CV_{\sigma^2})}$. A linear regression on CV_μ and $\sqrt{(CV_{\sigma^2})}$ could explain 93% of the variation in the source data. The regression has a standard error of 0.06. The relationship is influenced by the climatic severity index.

The regression line has the following equation:

$$CV_{\sigma^2} = (0.67CV_\mu - 0.04CSI + 0.47)^2 \qquad (117)$$

The relationship is strong and with a good estimate of CV_μ should give good estimates of $\sqrt{(CV_{\sigma^2})}$.

7.5 CORRELATION

7.5.1 Specification of E(ρ)

The expected value of the correlation coefficient in a community is not strongly related to any of the other parameters or one of the easily measured attributes that characterises a community. The correlation varies between 0.04 and 0.24. The average value can be taken as the most likely value and is equal to 0.13. The average value of the correlation coefficient is therefore specified as a fuzzy number.

More research needs to be done to understand what influences the correlation coefficient.

7.5.2 Specification of $\rho_{\delta\sigma}$ and $\rho_{\delta\sigma^2}$

No strong link between either $\rho_{\delta\sigma}$ or $\rho_{\delta\sigma^2}$ and any of the other load parameters or community attributes was found. These correlation coefficients are however strongly correlated and a linear regression with the following equality explains 86% of the variance in the source data.

$$\rho_{\delta\sigma^2} = \rho_{\delta\sigma} \tag{118}$$

A single fuzzy number is proposed for both correlation terms. The minimum and maximum values are calculated as the minimum and maximum values of both correlation terms measured in the source data. The minimum value is 0.61 and the maximum value is 0.95. The most likely value is 0.84.

7.6 SUMMARY AND CONCLUSION

The specification of load parameters for the voltage performance and loss calculation methods is discussed in this chapter. Some strong relationships between certain of the parameters have been identified and regression lines were fitted.

Weak relationships exist between the community attributes and the correlation parameters. For these relationships fuzzy numbers were proposed as a way of including their uncertainty into the calculation procedure.

The next chapter contains an illustration of how these parameters and relationships between parameters can be used to quantify the uncertainties in low voltage feeder design.

7.7 LOAD PARAMETERS USED IN REGRESSION ANALYSIS

This section contains a table with the source data used in the regression analysis. The household income was inflation adjusted to 1999 prices.

Table 24 Load parameters, household income, time with electricity and CSI for the townships used in the regression analyses

Name of township	Household income per month [Rand]	Time with electricity [Years]	CSI	$E(\mu)$	$E(\sigma)$	$E(\sigma^2)$	G	G_2
Claremont 96	3964	10	0	6.44	6.56	46.58	0.54	8.03
Claremont 97	3854	11	0	7.18	7.00	52.27	0.46	7.31
Claremont 98	3826	11	0	7.10	6.61	47.52	0.31	4.62
Cloetesville 94/5	4191	9	1	4.17	4.61	22.61	0.74	6.33
Helderberg 97	7000	4	1	5.62	6.70	46.67	0.61	8.08
Helderberg 98	7200	4	1	5.24	6.21	40.46	0.57	7.17
Helderberg 99	7465	5	1	5.45	6.20	39.97	0.48	6.03
Kwazakhele 95/6	1273	2	0.5	1.40	2.02	5.39	1.01	4.57
Lotus Park 99	2724	13	-1	4.30	4.61	22.11	0.60	5.31
Lotus Prk 98	2348	11	-1	4.52	4.82	24.25	0.59	6.66
Manyasteng 97	1144	5	0	1.80	2.21	8.64	0.88	6.12
Manyatseng 96	984	2	0	1.44	1.81	5.82	0.67	4.60
Orient Hills 98	1315	11	-1	3.11	3.61	14.65	0.86	6.00
Orient Hills 99	1844	13	-1	3.04	3.65	14.44	0.77	5.50
Rontree Estate 99	13304	14	1	7.11	6.98	53.16	0.63	9.55
Sanctuary Gdns 99	9667	1	0	3.87	5.62	33.05	0.82	9.12
Summerstrand 99	10759	1	0.5	3.96	4.78	25.97	0.70	6.83
Sweetwaters 96	1650	5	-0.5	1.04	1.58	3.91	0.93	5.52
Sweetwaters 97	1286	6	-0.5	1.45	1.87	5.95	0.71	4.01
Tafelsig 98	1985	8	1	2.76	3.53	13.83	0.75	5.92
Tafelsig 99	2678	9	1	2.75	3.47	13.33	0.76	5.95
Tambo 98	413	3	0	0.31	0.60	0.80	1.47	4.06
Umgaga 98	1143	3	-1	1.42	1.57	3.53	0.56	2.20
Umlazi 95/6	3020	6	-1	1.06	1.15	1.67	0.79	2.48
Umlazi 98	2274	7	-1	3.40	3.54	14.77	0.74	5.79
Umlazi 99	2120	9	-1	2.94	3.26	13.25	0.81	6.21
Walmer 98	959	3	0.5	0.98	1.77	3.74	1.04	4.13
Walmer Est 97	913	3	0.5	0.92	1.68	3.66	1.07	4.48

Table 25 Load parameters for the townships used in the regression analyses (continued)

Name of township	CV_{μ}	CV_{σ^2}	CV_{σ}	$CV_{Floorarea}$	$E(\rho)$	$\rho_{\delta\sigma}$	$\rho_{\delta\sigma^2}$
Claremont 96	0.47	0.79	0.29	0.28	0.18	0.87	0.84
Claremont 97	0.48	0.75	0.26	0.41	0.19	0.86	0.87
Claremont 98	0.65	0.80	0.30	0.38	0.18	0.72	0.70
Cloetesville 94/5	0.32	0.69	0.26	0.09	0.22	0.83	0.77
Helderberg 97	0.35	0.62	0.20	0.22	0.16	0.86	0.86
Helderberg 98	0.38	0.66	0.22	0.25	0.11	0.83	0.83
Helderberg 99	0.35	0.63	0.20	0.32	0.12	0.74	0.74
Kwazakhele 95/6	0.73	0.99	0.57	0.21	0.15	0.89	0.88
Lotus Park 99	0.29	0.62	0.21	0.15	0.12	0.78	0.77
Lotus Prk 98	0.31	0.68	0.21	0.22	0.12	0.83	0.84
Manyasteng 97	1.13	1.26	0.89	0.63	0.12	0.91	0.91
Manyatseng 96	1.43	1.36	0.88	0.72	0.13	0.87	0.88
Orient Hills 98	0.40	0.79	0.36	0.10	0.11	0.85	0.82
Orient Hills 99	0.36	0.73	0.29	0.11	0.12	0.81	0.79
Rontree Estate 99	0.41	0.80	0.31	0.39	0.15	0.87	0.83
Sanctuary Gdns 99	0.35	0.64	0.22	0.06	0.10	0.90	0.92
Summerstrand 99	0.55	0.78	0.37	0.22	0.07	0.86	0.94
Sweetwaters 96	1.13	1.32	0.76	0.37	0.24	0.91	0.95
Sweetwaters 97	1.25	1.24	0.84	0.41	0.10	0.81	0.78
Tafelsig 98	0.52	0.82	0.34	0.11	0.14	0.91	0.91
Tafelsig 99	0.50	0.82	0.33	0.28	0.13	0.90	0.90
Tambo 98	1.25	1.56	1.12	Unknown	0.04	0.84	0.81
Umgaga 98	0.80	1.07	0.66	0.47	0.08	0.61	0.62
Umlazi 95/6	0.66	1.08	0.52	0.20	0.21	0.92	0.89
Umlazi 98	0.53	0.90	0.43	0.14	0.15	0.89	0.88
Umlazi 99	0.62	0.99	0.50	0.49	0.10	0.91	0.88
Walmer 98	0.64	0.90	0.45	0.50	0.09	0.82	0.85
Walmer Est 97	0.82	1.01	0.55	0.50	0.08	0.87	0.92

8. PRACTICAL APPLICATION

8.1 INTRODUCTION

The methods described in this thesis can be applied in two different contexts:

- Initial planning
- Upgrade planning

Parameters for the former are estimated on the basis of little information that is known about a community. Some of the parameters of the latter might be known through data collected in a billing system (sales etc.)

The only difference between the two applications is how the load parameters are obtained. During initial planning some attributes of the consumers can be used to estimate the future load. Before an upgrade operation is initiated, energy sales information can be used as better estimates of the load.

8.2 LOAD PARAMETERS

Two load parameters needs to be specified:

- The estimated consumption per household
- Some estimate of the difference between consumers expressed as a coefficient of variation

Suppose a design of a low voltage feeder needs to be made to accommodate the first seven years of load growth.

In South Africa the average consumption per household after 7 years since electrification can be estimated using DTPet [17]. An estimate of the difference between consumers can be estimated with the knowledge of the difference in floor area of the houses in a community. This can be obtained from aerial photographs of a community.

Consider a community with an income of R3500 (in terms of 1999 prices) per household per month. The average consumption is estimated to be 1200 kWh. This relates to μ =7.28 A at 230 V.

Suppose an aerial survey of the community revealed that the coefficient of variation of the community is 0.5. The community is located in an area where the climatic severity index is zero.

The parameters for the voltage performance and loss calculation procedures can now be estimated using the equations in chapter 7. Table 26 contains the estimated parameters and the calculations are shown in section 8.7.

Table 26 Summary of calculated parameters for fuzzy probabilistic methods

Parameter	**Value**
μ	7.28
$V\delta$	13.66
$E\delta3$	42.64
$V\delta2$	528.84
$E[\sigma]$ or C_o	7.41
$E[\sigma^2]$ or C_{o2}	51.35
G	(0.146, 0.572, 0.647)
G2	(5.56,7.66, 8.66)
ρ	(0.04, 0.13, 0.24)

8.3 FEEDER TO BE EVALUATED

The following feeder will be evaluated:

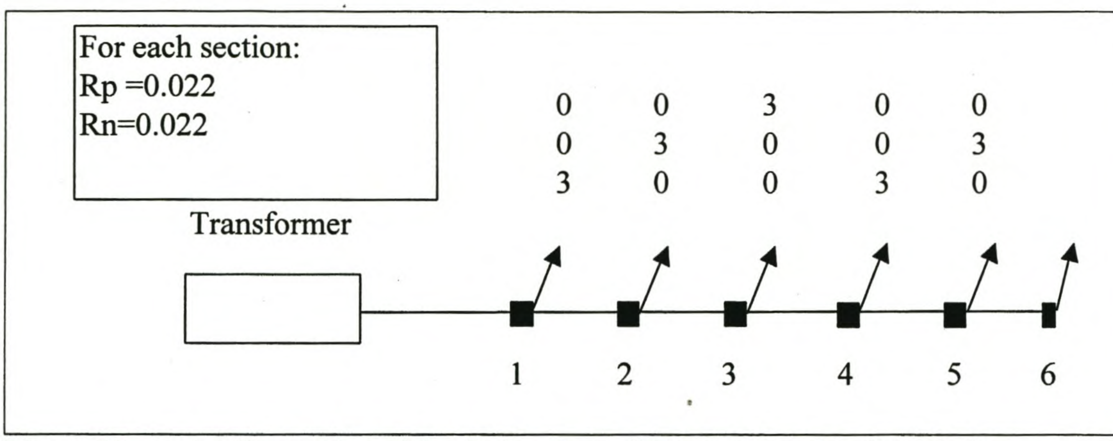

Figure 25 An illustration of the feeder that will be evaluated in this section

Table 27 Specification of proposed feeder

	Number of consumers			**Conductor resistance per section (Ω)**	
Node	**ma**	**mb**	**mc**	**Rp**	**Rn**
1			3	0.022	0.022
2		3		0.022	0.022
3	3			0.022	0.022
4			3	0.022	0.022
5		3		0.022	0.022
6	3			0.022	0.022

The conductor resistance and supply voltage uncertainty are taken as 0%.

8.4 FUZZY PROBABILISTIC VOLTAGE PERFORMANCE CALCULATIONS

The fuzzy probabilistic voltage performance can be calculated using the following steps:

Step 1: Specify a_k for each consumer on the feeder

The a-phase is the phase for which the voltage performance is calculated. If consumer i is connected to the a-phase on a three phase LV feeder, a_i is calculated as follows:

$$a_i = R_{pi} + R_{ni} \tag{119}$$

else

$$a_i = -0.5R_{ni} \tag{120}$$

where

R_{pi} is the phase resistance from the source to consumer i

R_{ni} is the neutral resistance from the source to consumer i

Table 28 Calculated values of a_k for the 18 consumers on the feeder

k	Ak	k	Ak	k	Ak
1	-0.011	7	0.132	13	-0.055
2	-0.011	8	0.132	14	-0.055
3	-0.011	9	0.132	15	-0.055
4	-0.022	10	-0.044	16	0.264
5	-0.022	11	-0.044	17	0.264
6	-0.022	12	-0.044	18	0.264

Step 2: Calculate K_1

$$K_1 := C_{o2} \cdot \sum_{j=0}^{N-1} (a_j)^2 + C_o^2 \cdot \rho \cdot \left(\sum_{k=0}^{N-1} a_k \right)^2 - C_o^2 \cdot \rho \cdot \left[\sum_{m=0}^{N-1} (a_m)^2 \right]$$

$$K_1 = 16.76721$$

where

N=18 and ρ=0.13

Step 3: Calculate the expected value of the percentile voltage drop

$$E(V\%tile) := V_s - \left(\mu \cdot \sum_{i=0}^{N-1} a_i + z_\alpha \cdot \sqrt{K_1} \right)$$

E(V%tile)= 212.138

where the confidence level is

$$cl = \frac{31*6*24-7}{31*6*24} = 0.998$$

and the ordinate for the standard normal distribution at this confidence level is

$z_\alpha = 2.954$

The supply voltage is taken as 230V

Step 4: Calculate the variance of the percentile voltage drop

$$V(V\%tile) := V\delta \sum_{k=0}^{N-1} \left(a_k\right)^2 \cdot \left(1 + \frac{a_k \cdot z_\alpha \cdot G2}{2 \cdot \sqrt{K_1}} + \frac{z_\alpha \cdot GC_0 \cdot \rho}{2 \cdot \sqrt{K_1}} \cdot \sum a - \frac{z_\alpha \cdot GC_0 \cdot \rho}{2 \cdot \sqrt{K_1}} \cdot a_k \right)^2$$

V(V%tile)= 11.44995

The most likely values of G, G2 and ρ, are used in the example calculation.

Step 5: Calculate the minimum and maximum percenitle voltage drops

$$\text{Maximum} = V_s - \sum_{a_i<0} a_i C_k$$

$$\text{Minimum} = V_s - \sum_{a_i>0} a_i C_k$$

$$\text{Maximum} = 253.76 \text{ V}$$

$$\text{Minimum} = 158.72 \text{ V}$$

where

C_k is the circuit breaker size or maximum value of the load current for consumer k, for the example taken to be 60

V_s is the supply voltage (230 V)

Step 6: Calculate the scaled values of E(%tile) and V(%tile), μ_{ns} and σ^2_{ns}

$$\mu_{ns} := \frac{(E(V\%tile) - Minimum)}{Maximum - Minimum}$$

$$\mu_{ns} = 0.562$$

$$\sigma_{ns} := \left[\frac{\sqrt{V(V\%tile)}}{(Maximum - Minimum)}\right]$$

$$\sigma_{ns}^{2} = 1.268 \times 10^{-3}$$

Step 7: Calculate α and β, the beta parameters of the distribution of voltage performance percentile values

$$\alpha := \frac{\left[\mu_{ns}^{2} - \left[\left(\sigma_{ns}^{2} + \mu_{ns}^{2}\right)\cdot\mu_{ns}\right]\right]}{\sigma_{ns}^{2}}$$

$$\beta := \frac{\alpha\left(1 - \mu_{ns}\right)}{\mu_{ns}}$$

$$\alpha = 108.579$$

$$\beta = 84.602$$

Step 8: Calculate the most likely design voltage value

$$V_{design} := qbeta(dcl, \alpha, \beta)\cdot(Maximum - Minimum) + Minimum$$

$$V_{design} = 207.781$$

where

qbeta is the inverse beta distribution function

dcl is the design confidence level taken as 10% (see discussion below)

Step 9: Calculate the minimum and maximum values of the design voltage

An exhaustive search of the problem space was performed to determine the fuzzy number for V_{design}.

The following fuzzy number for V_{design} was obtained:

$V_{design} = (207.14, 207.78, 209)$

Step 10: Calculate a single design voltage value

In order to make a design decision, a single design value is required. The design value can be obtained at some level of presumption. If a level of presumption of 0.5 is taken, the design value is:

$$V_{\alpha=0.5} = V_{likely} - (V_{likely} - V_{min})0.5 = 207.46$$

Note that obtaining the design value in this manner assumes that the possibility distribution is triangular. Since all parameter uncertainties are modelled as triangular fuzzy numbers and the calculation procedure is non-linear, the design value obtained in this manner is approximate with reasonable accuracy. The result is however readily obtained and is considered acceptable for design purposes. More accurate values can be achieved by applying non-linear programming at different α-level cuts of the parameter uncertainty fuzzy numbers.

8.4.1 Choice of the design confidence level and the level of presumption

The choice of the design confidence level is a business decision since a lower design confidence level exposes the owner of the designed feeder. The owner could play-off additional risk against potential financial gain achieved by delaying capital investment. The design confidence level would also be different for feeders designed with and without an upgrade path.

Choice of level of presumption is made based on the certainty that is required in a decision. If it is very important that the result is correct a higher level of presumption is used.

The choice of the values to be used for the confidence and presumption levels is a business decision. At the time of writing, work was being undertaken by the author on behalf of Eskom to identify the most appropriate values for electrification in South Africa. Completely different values could be appropriate for electrification or refurbishment in other countries.

It should be noted that prior to the choice of these values the use of the methods described in this text has limited application. This because the design value is obtained by specifying the confidence level and level of presumption. It is very important to understand the consequence of obtaining a design value at some level of confidence and also at a level of presumption.

8.5 FUZZY PROBABILISTIC LOSS CALCULATIONS

In a similar fashion a fuzzy-probabilistic assessment of the resistive loss in the sample feeder may be calculated.

Step 1 : Define Rp, Rn, Qa, Qb and Qc

Rp and Rn are vectors of the phase and neutral resistances for different sections. Element Rp_i is the phase resistance between node i and i+1.

$$Rp = \begin{pmatrix} 0.022 \\ 0.022 \\ 0.022 \\ 0.022 \\ 0.022 \\ 0.022 \end{pmatrix}$$

$$Rn = Rp$$

Table 29 contains matrices with the elements of Qa, Qb and Qc with a column for each section and a row for each consumer. Element $Qa_{i,j}$ is 1 if load current from consumer i is present in section j.

Table 29 Calculated values for Qa, Qb and Qc

Qa	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	1	1	1	1	1	1
4	1	1	1	0	0	0
5	1	1	1	0	0	0
6	1	1	1	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	0	0	0	0	0	0

Qb	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	1	1	1	1	1	0
8	1	1	1	1	1	0
9	1	1	1	1	1	0
10	1	1	0	0	0	0
11	1	1	0	0	0	0
12	1	1	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	0	0	0	0	0	0

Qc	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0
13	1	1	1	1	0	0
14	1	1	1	1	0	0
15	1	1	1	1	0	0
16	1	0	0	0	0	0
17	1	0	0	0	0	0
18	1	0	0	0	0	0

Step 2 : Calculate $E[K1]$, $E[K3_{ij}]$, $E[J_k]$, $E[P_{ij}]$, K_2, $K4_{ij}$, $D[J_k(0)]$, $D[P_{ij}(0)]$, $D2[J_k(0)]$

Calculate the most likely value of the constants used the remainder of the loss calculations:

$$E[K1] = \mu^2 + C_{o2}$$

$$E[K1] = 104.348$$

$$E[K3\_ij] = \mu^2 + \rho \cdot C_o^2$$

$$E[K3\_ij] = 60.126$$

$$E[J\_k] = E[K1] + V\delta$$

$$E[J\_k] = 118.003$$

$$E[P\_ij] = E[K3\_ij]$$

$$K2 = 2\mu + G2$$

$$K2 = 22.22$$

$$K4\_ij = \mu + \rho \cdot G \cdot C_o$$

$$K4\_ij = 7.831$$

$$D[J\_k(0)] = K2$$

$$D[P\_ij(0)] = K4\_ij$$

$$D2[J\_k(0)] = 2$$

Step 3 : Calculate the most likely value of the expected loss, E[Loss]

$$E[Loss1] = \sum_{s=0}^{Sections-1} \left[\sum_{k=0}^{N-1} E[J\_k] \left(Qa_{k,s} + Qb_{k,s} + Qc_{k,s} \cdot \left(Rp_s + Rn_s \right) \right] \right.$$

$$E[Loss1] = 327.104$$

$$E[Loss2] = \sum_{s=0}^{Sections-1} \left[\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Qa_{i,s} \cdot Qa_{j,s} + Qb_{i,s} \cdot Qb_{j,s} + Qc_{i,s} \cdot Qc_{j,s} \cdot \left(Rn_s + Rp_s \right) \right] \right.$$

$$E[Loss2] = 619.062$$

$$E[Loss3] = \sum_{s=0}^{Sections-1} \left[-\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i \neq j) \cdot \left[E[P\_ij] \left(Qa_{i,s} \cdot Qb_{j,s} + Qa_{i,s} \cdot Qc_{j,s} + Qb_{i,s} \cdot Qc_{j,s} \cdot Rn_s \right] \right.$$

$$E[Loss3] = -345.246$$

$$E[Loss] = E[Loss1] + E[Loss2] + E[Loss3]$$

$$E[Loss] = 600.92 \ \mathrm{W}$$

Step 4 : Calculate most likely value of $T_1(k)$ and $T_2(k)$ for each consumer k

$$T_{1a}(k) = \sum_{s=0}^{Sections-1} \left[D[J\_k(0)] \left(Qa_{k,s} + Qb_{k,s} + Qc_{k,s} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$$T_{1b}(k) = \sum_{s=0}^{Sections-1} 2 \left[\sum_{i=0}^{N-1} (i \neq k) \cdot D[P\_ij(0)] \left(Qa_{k,s} \cdot Qa_{i,s} + Qb_{k,s} \cdot Qb_{i,s} + Qc_{k,s} \cdot Qc_{i,s} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

$$T_{1c}(k) = \sum_{s=0}^{Sections-1} - \left[\sum_{i=0}^{N-1} D[P\_ij(0)] \left(Qa_{k,s} \cdot Qb_{i,s} + Qb_{k,s} \cdot Qc_{i,s} + Qa_{k,s} \cdot Qc_{i,s} + Qa_{i,s} \cdot Qb_{k,s} + Qb_{i,s} \cdot Qc_{k,s} + Qa_{i,s} \cdot Qc_{k,s} \right) \cdot \left(Rn_s \right) \right]$$

$$T_1(k) = T_{1a}(k) + T_{1b}(k) + T_{1c}(k)$$

$$T_2(k) = \sum_{s=0}^{Sections-1} \left[D2[J\_k(0)] \left(Qa_{k,s} + Qb_{k,s} + Qc_{k,s} \right) \cdot \left(Rp_s + Rn_s \right) \right]$$

Table 30 Calculated values for T1 and T2 for each of the consumers

k	T1	T2
1	14.14	0.53
2	14.14	0.53
3	14.14	0.53
4	8.62	0.26
5	8.62	0.26
6	8.62	0.26
7	9.20	0.44
8	9.20	0.44
9	9.20	0.44
10	5.23	0.18
11	5.23	0.18
12	5.23	0.18
13	4.77	0.35
14	4.77	0.35
15	4.77	0.35
16	2.36	0.09
17	2.36	0.09
18	2.36	0.09

Step 5 : Calculate the most likely value of the variance of the loss

$$V[Loss] = \sum_{i=0}^{N-1} \left(V\delta \cdot T_1(i)^2 + V\delta 2 \frac{T_2(i)^2}{4} + E\delta 3 \cdot T_1(i) \cdot T_2(i) \right)$$

$$V[Loss] = 1.937 \times 10^4$$

Step 6 : Calculate the maximum values of the loss

This maximum value is the absolute maximum loss used in fitting the beta parameters and uses the circuit breaker size, c_k, as the maximum current. The minimum and maximum design loss values are calculated in step 9.

$$Max_1 = \sum_{s=0}^{Sections-1} \left(\sum_{k=0}^{N-1} c_k \cdot Qa_{k,s} \right)^2 \cdot \left(Rn_s + Rp_s \right)$$

$$Max_1 = 2.138 \times 10^4$$

$$Max_2 = \sum_{s=0}^{Sections-1} \left(\sum_{k=0}^{N-1} c_k \cdot Qb_{k,s} \right)^2 \cdot \left(Rn_s + Rp_s \right)$$

$$Max_2 = 1.568 \times 10^4$$

$$Max_3 = \sum_{s=0}^{Sections-1} \left(\sum_{k=0}^{N-1} c_k \cdot Qc_{k,s} \right)^2 \cdot \left(Rn_s + Rp_s \right)$$

$$Max_3 = 9.979 \times 10^3$$

$$Maximum = Max_1 + Max_2 + Max_3$$

$$Maximum = 4.704 \times 10^4$$

Step 7 : Calculate α and β, the parameters of a beta distribution for the most likely loss distribution

$$\mu_{ns} = \frac{E[Loss]}{Maximum} \qquad \sigma_{ns} = \frac{\sqrt{V[Loss]}}{Maximum}$$

$$\mu_{ns} = 0.013 \qquad {\sigma_{ns}}^2 = 8.754\times 10^{-6}$$

$$\alpha = \frac{\left[{\mu_{ns}}^2 - \left[\left({\sigma_{ns}}^2 + {\mu_{ns}}^2\right)\cdot\mu_{ns}\right]\right]}{{\sigma_{ns}}^2} \qquad \beta = \frac{\alpha\left(1-\mu_{ns}\right)}{\mu_{ns}}$$

$$\alpha = 18.387 \qquad \beta = 1.421\times 10^3$$

A beta distribution with these parameters is very skew and further confirms the appropriateness of the beta distribution. A normal distribution would not adequately represent the distribution of the resistive loss.

Step 8 : Calculate the most likely loss design percentile

$$Loss_{design} = qbeta(dcl, \alpha, \beta)\cdot Maximum$$

$$Loss_{design} = 784.775$$

qbeta is the inverse beta probability distribution with parameters α and β

dcl is the design percentile value, in this case dcl = 0.9

Step 9 : Calculate the minimum and maximum loss design percentile

The minimum and maximum values are calculated by repeating steps 2-8 with the minimum and maximum values for G, G2 and ρ.

The loss fuzzy design percentile is equal to (749.8;784.78;804.56)

Step 10: Calculate a single design loss value

A single design loss value can be extracted from the loss fuzzy design percentile by applying a presumption level or α-level. The single design value can be calculated as:

$$\text{Loss}_{\alpha=0.5} = \text{Loss}_{\text{likely}} + (\text{Loss}_{\text{maximum}} - \text{Loss}_{\text{likely}})0.5 = 794.67$$

Note that obtaining the design value in this manner assumes that the possibility distribution is triangular. Since all parameter uncertainties are modelled as triangular fuzzy numbers and the calculation procedure is non-linear, the design value obtained in this manner is approximate with reasonable accuracy. The result is however readily obtained and is considered acceptable for design purposes. More accurate values can be achieved by applying non-linear programming at different α-level cuts of the parameter uncertainty fuzzy numbers.

8.5.1 Choice of the design confidence level and level of presumption

The distribution of loss can be used in life cycle cost calculation of a feeder. If a number of feeders in a system are evaluated then the distribution of losses can be added to obtain a total loss for the network. As feeders are added, the distribution of the total loss will be Gaussian due to the central limit theorem. For a very large number of feeder the design percentile will approach the mean value and the specification of the design confidence level is not very significant.

If the loss calculations are made to minimize the life cycle cost of a feeder, then the specification of the design confidence level is a business decision (see section 8.4.1) that should be taken with the choice of the confidence level of the voltage performance design.

8.6 SUMMARY AND CONCLUSIONS

This chapter shows by way of example, step by step how to use the methods described in previous chapters. All the calculation, values of input variables, calculated constants and results are shown.

The voltage performance calculation and loss calculations are made after specifying the average of the load current means and the standard deviation of the load current means. Some of the uncertainties in the input parameters are treated as fuzzy numbers and the results of the calculations are fuzzy numbers. Unique design values are obtained from the fuzzy numbers by applying a level of presumption.

8.7 APPENDIX: CALCULATIONS OF LOAD PARAMETERS

Step 1 : Calculation of Vδ

$$\mu = 7.28 \qquad CSI = 0 \qquad CV_{FloorArea} = 0.5$$

$$CV_{\mu} = CV_{FloorArea} - 0.08 \cdot \mu + 0.59$$

$$V\delta = \left(CV_{\mu} \cdot \mu\right)^2$$

$$V\delta = 13.655$$

Step 2 : Calculation of Eδ3, the third moment of δ and Vδ2, the variance of δ²

$$\mu_{ns} = \frac{\mu}{c_k} \qquad \sigma_{ns} = \frac{CV_{\mu} \cdot \mu}{c_k}$$

$$\alpha = \frac{\left[\mu_{ns}^{2} - \left[\left(\sigma_{ns}^{2} + \mu_{ns}^{2}\right) \cdot \mu_{ns}\right]\right]}{\sigma_{ns}^{2}} \qquad \beta = \frac{\alpha\left(1 - \mu_{ns}\right)}{\mu_{ns}}$$

$$\alpha = 3.289 \qquad \beta = 23.817$$

$$E\mu 3 = \frac{\alpha \cdot (\alpha + 1) \cdot (\alpha + 2)}{(\alpha + \beta)(\alpha + \beta + 1) \cdot (\alpha + \beta + 2)} \cdot c_k^{3}$$

$$E\mu 4 = \frac{\alpha \cdot (\alpha + 1) \cdot (\alpha + 2) \cdot (\alpha + 3)}{(\alpha + \beta)(\alpha + \beta + 1) \cdot (\alpha + \beta + 2)(\alpha + \beta + 3)} \cdot c_k^{4}$$

$$E\delta 3 = E\mu 3 - 3 \cdot V\delta \cdot \mu - \mu^{3}$$

$$E\delta 4 = E\mu 4 - \mu^{4} - 6 \cdot \mu^{2} V\delta - 4\mu \cdot E\delta 3$$

$$V\delta 2 = E\delta 4 - V\delta^{2}$$

$$E\delta 3 = 42.637 \qquad V\delta 2 = 528.841$$

Step 3 : Calculation of Co and Co2

Note the Co = E[σ] and Co2 = E[σ²]

$$Co = 0.91 \cdot \mu + 0.25 \cdot CSI + 0.78$$

$$Co2 = 7.89 \cdot \mu + 2.64 \cdot CSI - 6.09$$

$$Co = 7.405$$

$$Co2 = 51.349$$

Step 4 : Calculation of V_σ and V_{σ^2}

$$CV_\sigma = 0.67 \cdot CV_\mu - 0.03 \cdot CSI$$

$$CV_{\sigma 2} = \left(0.67 \cdot CV_\mu - 0.04 \cdot CSI + 0.47\right)^2$$

$$CV_\sigma = 0.34$$

$$CV_{\sigma 2} = 0.656$$

$$V_\sigma = \left(CV_\sigma \cdot Co\right)^2$$

$$V_{\sigma 2} = \left(CV_{\sigma 2} \cdot Co2\right)^2$$

$$V_\sigma = 6.342$$

$$V_{\sigma 2} = 1.136 \times 10^3$$

Step 5 : Calculations of G and G2

The calculation is repeated for $\rho_{\delta\sigma}$=0.61 ; $\rho_{\delta\sigma}$=0.84 ; $\rho_{\delta\sigma}$=0.95. These are the minimum, most likely and maximum values for $\rho_{\delta\sigma}$.

$$\rho_{\delta\sigma 2} = 0.61$$

$$\rho_{\delta\sigma} = \rho_{\delta\sigma 2}$$

$$G = \rho_{\delta\sigma} \cdot \sqrt{\frac{V_\sigma}{V\delta}}$$

$$G2 = \rho_{\delta\sigma 2} \cdot \sqrt{\frac{V_{\sigma 2}}{V\delta}}$$

$$G = 0.416$$

$$G2 = 5.563$$

9. SUMMARY AND CONCLUSION

This section contains a summary of the work per chapter as presented in this thesis. The need for more research in specific areas are identified and summarized in section 9.2.

9.1 SUMMARY AND CONCLUSIONS PER CHAPTER

9.1.1 Chapter 2: Overview of uncertainties and quantification of uncertainties

Quantification techniques found in literature are reported in this chapter. A specific type of information is required in order to use some of the uncertainty models. For instance, a probabilistic model requires knowledge of the probability of different events occurring.

Probabilistic information is available for residential consumer loads but only possibilistic type information is available for the load parameter and conductor resistance uncertainties. A combined fuzzy-probabilistic model is proposed as the most suitable method for representing the uncertainty due to residential loads, their parameters and the specification of the network.

9.1.2 Chapter 3: Residential consumer load model

A model for residential consumer load is proposed in this chapter. Each individual load current is modelled as a beta probability distribution with a mean and standard deviation. It is assumed that the consumer loads are constant current in nature.

The means and standard deviations of the load current of consumers in a community have distributions, each with a mean and a standard deviation. The distributions of the means of the load currents and standard deviations of load currents are correlated. This correlation can be used to fit a regression line in order to account for the covariance between the two distributions.

The residential load current uncertainty can then be represented with two random variables:

- The consumer mean load current
- The error in the regression between the means and standard deviations

The chapter further describes methods to calculate the distribution of a percentile value of the distribution of a linear combination of load currents. These methods form the basis for the voltage performance calculation and the resistive loss calculation methods described in subsequent chapters.

9.1.3 Chapter 4: Voltage performance calculation

A new voltage performance calculation method is described in this section and uses the load current described in chapter 3. The method can be significantly simplified by using an assumption that the percentile values of the distribution can be calculated using a normal distribution.

It is further proposed that the South African voltage performance standard (NRS048) can be estimated as a percentile of the voltage performance distribution. Using these assumptions, the NRS 048 voltage performance of a proposed feeder can be estimated. The method is probabilistic in nature and can be used to calculate the probability that a feeder will comply with the NRS 048 voltage performance criteria for a specified period of time. The method was tested with load data from low income and medium income communities.

9.1.4 Chapter 5: Resistive loss calculation

A new probabilistic method that was developed to calculate the resistive loss in a feeder is described in this chapter. The method uses the load model described in chapter 3 to calculate the resistive loss for a period of time. The resistive loss for a feeder has a range of probable values with a mean and standard deviation. A single design value can be obtained using a beta distribution.

The method was tested with load data from a low and medium income community and the results are suitable for practical design.

9.1.5 Chapter 6: Fuzzy probabilistic calculation methods

The load parameter and network parameter uncertainties cannot be probabilistically described since no probabilistic information is available. The probabilistic methods described in chapter 4 and 5 are enhanced using fuzzy-probabilistic techniques.

The calculation of the fuzzy-probabilistic voltage performance distribution can be done using:

- Non-linear programming
- Exhaustive problem space search
- Genetic algorithms

The advantage of genetic algorithms is that a solution is readily obtained and although this solution might not be absolutely correct, it is close to the exact solution.

9.1.6 Chapter 7: Load parameter prediction

An analysis of the load data from the NRS and TSI load research projects is described in this chapter. Strong relationships between some of the load parameters for the voltage performance and loss calculation methods are identified. The result of the analysis is that the load parameters can be estimated using knowledge of the distribution of the consumption of different consumers in a community.

This makes the procedures ideal for upgrading scenarios where the consumption of the community is readily available. The initial designs can be revised before upgrading the network to make use of the additional information contained in the sales data.

9.1.7 Chapter 8: Practical application

In this chapter, the use of the design procedures is described in a step by step fashion. All calculations are shown with typical input values and results.

9.2 FUTURE RESEARCH

A number of topics for future research have been identified and are listed below:

1. The most appropriate modelling of residential loads, i.e. constant current, constant power or constant resistance.

2. Investigation of the effect of MV/LV transformer impedances on the accuracy of the calculation methods.

3. Expanding the methods described in this thesis to include the effect of temperature rise on the conductor resistance due to the variations in the load currents and ambient temperature.

4. The specification of cl, the confidence level (see section 4.5) to estimate the voltage performance of a feeder with respect to EN50160, the euro-norm voltage performance criterion. The assumption about the distribution of periods of high load as a function of the number of days with voltage performance violations could also be investigated.

5. The calibration of the confidence level in the Herman Beta method, to reflect the probability that a designed feeder will not meet the NRS 048 voltage performance criterion. This will allow users of the Herman Beta method to benefit from this research.

6. An assumption of triangular membership functions for load and network parameter uncertainty was used. The impact of this choice, which was made mainly for simplicity, should be investigated.

7. The choice of genetic algorithm parameters used in the fuzzy-probabilistic method. The aim would be to find a set of parameters which will ensure that the methods perform optimally, i.e. consistently give accurate results within an acceptable time.

8. The improvement of the estimating procedures for the load parameters, especially the correlation between the loads of individual residential consumers should be investigated. An examination into improving the robustness of the fitted models and the use of non-linear and other non-parametric techniques could be included in this study.

10. REFERENCES

1. RN Allan, B Borkowska, CH Grigg, "Probabilistic analysis of power flows", Proceedings of the IEE , Vol 121, No 12, December 1974
2. GJ Anders, "Probabilistic concepts in electric power systems", John Wiley & Sons, ISBN 0-471-50229-4, 1990
3. DB Belzer; MA Kellogg, "Incorporating sources of uncertainty in forecasting peak power loads - A Monte Carlo analysis using the extreme value distribution", IEEE Transactions on Power Systems, Vol 8 No2, p 730-737, May 1993
4. R Billinton; W Li, "A Monte Carlo method for multi-area generation system reliability assessment", IEEE Transaction on Power Systems, Vol 7 No 4 p1487-1492, Nov 1992
5. B Borkowska, "Probabilistic Load Flow", IEEE PES Summer Meeting and EHV/UHV Conference, T 73 480-0
6. M Brucoli, F Torelli, R Napoli, "Quadratic probabilistic load flow with linearly modeled dispatch", Electric Power and Energy Systems, Vol 7 No3 July 1985
7. Arthur M. Breipohl; Fred N. Lee; Duanran Zhai; Rambabu Adapa , "A Gauss-Markov load model for application in risk evaluation and production simulation.", IEEE Transactions on Power Systems v 7 n 4 Nov 1992 p 1493-1499
8. Robert P. Broadwater; Hesham E. Shaalan; Wolter J. Fabrycky; Robert E. Lee , "Decision evaluation with interval mathematics: a power distribution system case study", IEEE Transactions on Power Delivery v 9 n 1 Jan 1994. p 59-67
9. William J. Burke; Hyde M. Merrill; Fred C. Schweppe; Bryne E. Lovell; Michael F. McCoy; Sharron A. Monohon , "Trade off methods in system planning.", IEEE Transactions on Power Systems v 3 n 3 Aug 1988 p 1284-1290
10. Che-Hao Chang; Yeou-Koung Tung; Jinn-Chuang Yang , "Monte Carlo simulation for correlated variables with marginal distributions", Journal of Hydraulic Engineering v 120 n 3 Mar 1994. p 313-331

11. Che-Hao Chang; Jinn-Chuang Yang; Yeou-Koung Tung , "Uncertainty analysis by point estimate methods incorporating marginal distributions", Journal of Hydraulic Engineering v 123 n 3 Mar 1997. p 244-250

12. J-F Chen; W-M Wang; C-M Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting", Electric Power System Research, Vol 34 p187-196, 1995

13. Chao-shun Chen; Tsung-Hsien Wu; Chung-Chieg Lee; Yenn-Minn Tzeng, "The application of load models of electric appliances to distribution system analysis", IEEE Transactions of Power Systems, Vol 10, No 3, August 1995

14. Enrique O. Crousillat; Peter Dorfner; Pablo Alvarado; Hyde M. Merrill , "Conflicting objectives and risk in power system planning", IEEE Transactions on Power Systems v 8 n 3 Aug 1993. p 887-895

15. Davis, L; "Handbook of genetic algorithms", ISBN 0-442-00173-8, Library of congress

16. M Dekenah, S Heunis; "Analysis of NRS LR project, Load and socio demographic data: 2000", ref 20000605SH, NRS LR Project, Contact M Dekenah 012 667 1347

17. Marcus Dekenah; SW Heunis, "Pre-electrification tool 99", Electrification prediction tool developed for Eskom DT group by Marcus Dekenah Consulting cc, tel 012 667 1347

18. SB Dhar, "Power system long range decision analysis under fuzzy environment", IEEE Transactions on Power apparatus and systems, Vol PAS-98 No 2 p585-596, April 1979

19. A. D. Dimitrovski, "Iterative Fuzzy Load Flow in Radial Distribution Networks", PMAPS 2000, 6 th International conference on probabilistic methods applied to power systems, PSP2-118, ISBN 972-95194-2-0

20. D Dubois; H Prade, Fuzzy sets and systems: Theory and application, Academic press 1980

21. B Efron; RJ Tibshirani, "An introduction to the bootstrap", Chapman & Hall 1993

22. "Voltage characteristics of electricity supplied by public distribution systems", EN 50160:2000

23. S. A. Farghal; R. M. El-Dewieny; M. Roshdy Abdel Aziz , "Generation expansion planning using the decision tree technique.", Electric Power Systems Research v 13 n 1 Aug 1987 p 59-70

24. Timothy K. Gates; Muhammad A. Al-Zahrani , "Spatiotemporal stochastic open-channel flow. I: Model and its parameter data", Journal of Hydraulic Engineering v 122 n 11 Nov 1996. p 641-650

25. CT Gaunt; R Herman; M Dekenah; RL Sellick; SW Heunis, "Data collection, load modelling and probabilistic analysis for lv domestic electrification", CIRED 1999, session 5, paper 20

26. CT Gaunt, "Beta parameters for residential loads", The transactions of the SA institute of electrical engineers, Vol 90, no 1, March 1999, p35

27. J Goslin, Xgen, autockr@ix.netcom.com

28. X Guan; W.H.E. Liu; A. D. Papalexopoulos, "Application of fuzzy set method in an optimal power flow", Electric Power System Research 34 (1995) 11-18

29. R Herman, "Voltage regulation analysis for the design of low voltage networks feeding stochastic domestic electrical loads", PhD Thesis, University of Stellenbosch, South Africa

30. R Herman; SW Heunis, "General probabilistic voltage drop calculation method for LV distribution networks based on a beta pdf load model", Electric Power Systems Research Vol 46 p 45-49, 1998

31. R Herman; JJ Kritzinger, "The statistical description of grouped domestic electrical load currents", Electric Power Systems Research, 27 (1993) 43-48

32. R Herman, "Using the bootstrap method for evaluating variations due to sampling in Voltage Regulation Calculations", Proceedings of SAUPEC 1996, p 235-238

33. R Herman; JS Maritz; JHR Enslin, "The analysis of voltage regulation in residential distribution networks using the beta distribution model", Electric Power Systems Research 29 (1994) 213-216

34. IEEE task force on load representation for dynamic performance, system dynamic performance sub committee, power system engineering committee, "Bibliography on load models for power flow and dynamic performance simulation", IEEE/PES 1994, Summer Meeting, San Francisco, 94 SM 580-1 PWRS

35. “Standard voltages”, IEC 38

36. Peggy A. Johnson; Bilal M. Ayyub , "Modelling uncertainty in prediction of pier scour", Journal of Hydraulic Engineering v 122 n 2 Feb 1996. p 66-72

37. N. Kagan; R.N. Adams , "Electrical power distribution systems planning using fuzzy mathematical programming", International Journal of Electrical Power and Energy System v 16 n 3 Jun 1994. p 191-196

38. GK Kanji, "100 Statistical tests", Sage publications, ISBN 0-7619-6151-8

39. OA Klitin, "Stochastic load flows", IEEE Transactions on Power Apparatus and Systems, Vol, PAS-94, no.2 , March/April 1975

40. Leuenberger, DG, "Linear and nonlinear programming", Addison Wesley publishing company, ISBN 0-201-15794-2

41. Jeeng-Min Ling; Chin-E. Lin; Ching-Lien Huang , "Investigation of fuzzy real power flow modelling with probabilistic-heuristic based information", Electric Power Systems Research v 32 n 3 Mar 1995. p 175-186

42. Jeeng-Min Ling; Chin-E. Lin; Ching-Lien Huang , "Probability measure of adequacy assessment using a fuzzy approach", Electric Power Systems Research v 33 n 1 Apr 1995. p 7-15

43. Jeeng-Min Ling; Chin-E. Lin; Ching-Lien Huang , "Possibility measures of transmission capacity assessment using a fuzzy approach", Electric Power Systems Research v 33 n 3 Jun 1995. p 165-173

44. E Lloyd, Handbook of applicable Mathematics, Volume II Probability, J Wiley & sons

45. C-H Loh; W-Y Jean, "Uniform-hazard response spectra - an alternative approach", Earthquake engineering and structural dynamics, Vol 23, 433-445, 1994

46. Hyde M. Merrill; Allen J. Wood , "Risk and uncertainty in power system planning.", Electrical Power & Energy Systems v 13 n 2 Apr 1991 p 81-90

47. M Mitchel, An introduction to genetic algortihms, Massachusetts Instutue of Technology, ISBN 0-262-13316-4

48. V. Miranda; J. T. Saraiva , "Fuzzy modelling of power system optimal load flow.", 91 IEEE Power Ind Comput Appl Conf Presented 17 PICA Conf. Publ by IEEE, IEEE Service Center, Piscataway, NJ, USA (IEEE cat n 92CH2948-8). p 386-392

49. J.A. Momoh; K. Tomsovic , "Overview and literature survey of fuzzy set theory in power systems", IEEE Transactions on Power Systems v 10 n 3 Aug 1995. p 1676-1690

50. Robert J. Mulvihill , "Design-safety enhancement through the use of hazard and risk analysis.", IEEE Transactions on Reliability v 37 n 2 Jun 1988 p 149-158

51. NRS 034-1:1997 Rationalised User Specification, Electricity Distribution - Guidelines for the provision of Electrical Distribution Networks in Residential Areas, ISBN 0-626-11235-4 ,1997

52. "NRS 048-2:1996 ELECTRICITY SUPPLY - QUALITY OF SUPPLY", ISBN 0-626-11069-6, 1996

53. NRS Load research database, project manager Marcus Dekenah of Marcus Dekenah Consulting cc. 012 667 1347, mdekenah@pixie.co.za

54. Simon Olafsson; Ragnar Sigbjornsson , "Application of ARMA models to estimate earthquake ground motion and structural response", Earthquake Engineering & Structural Dynamics v 24 n 7 Jul 1995. p 951-966

55. J. Tome Saraiva; Vladimiro Miranda; L.M.V.G. Pinto , "Impact on some planning decisions from a fuzzy modelling of power systems", IEEE Transactions on Power Systems v 9 n 2 May 1994. p 819-825

56. R.J. Sarfi; M.M.A. Salama; A.Y. Chikhani , "Applications of fuzzy sets theory in power systems planning and operation: A critical review to assist in implementation", Electric Power Systems Research v 39 n 2 Nov 1996. p 89-101

57. HR Sirisena, EPM Brown, "Representation of non-Gaussian probability distributions in stochastic load-flow studies by the method of Gaussian sum approximations", IEE Proceedings Vol 130 part C, No 4 July 1983

58. Dipti Srinivasan; A.C. Liew; C.S. Chang , "Applications of fuzzy systems in power systems", Electric Power Systems Research v 35 n 1 Oct 1995. p 39-43

59. R.H. Stillman; M.S. Mackisack; M. Griffin , "Modelling the reliability of distribution line structures", Electric Power Systems Research v 41 n 2 May 1997. p 85-90

60. Dusan Teodorovic; Milica Kalic , "Fuzzy route choice model for air transportation networks", Transportation Planning and Technology v 19 n 2 1995. p 109-119

61. TSI load research project, Managed by Vic Lawrence +27 11 629 5111, Vic.Lawrence@Eskom.co.za

62. Y-K Tung; LW Mays, "Risk analysis for hydraulic design", Journal of the hydraulics division, p 893-913, May 1980

63. Guang-Yuan Wang; Jin-Ping Ou , "Fuzzy random vibration of multi-degree-of-freedom hysteretic systems subjected to earthquakes.", Earthquake Engineering & Structural Dynamics v 15 n 5 Jul 1987 p 539-548

64. H. Lee Willis; R.W. Powell; H.N. Tram;, "Long range distribution planning with load forecast uncertainty", IEEE Transactions on Power systems; Vol 2 No 3 p 684-691, Aug 1987

65. LA Zadeh, "Fuzzy sets", Information and control, pp 338-353 August 1965.

66. D Zhai; A.M. Breipohl; F.N. Lee; R. Adapa, "The effect of load uncertainty on unit commitment risk", IEEE transactions on power systems, Vol 9 No1 p510-517, February 1994